An Experimental Approach to Multiple Case Constructions in Korean

Yong-hun Lee*

Chungnam National University &
Hannam University

Yong-hun Lee. 2013. An Experimental Approach to Multiple Case Constructions in Korean. *Language and Information 17.2*, 29–50. Multiple Nominative Constructions (MNCs) and Multiple Accusative Constructions (MACs) have been some of the hottest and interesting topics in Korean syntax. This paper took empirical approaches to these constructions and examined native speakers' grammaticality judgements of these constructions. Though there are lots of previous studies on these constructions, Ryu (2010, 2013a, 2013b, 2013c) recently tried to unify MNCs and MACs into Multiple Case Constructions (MCCs) and to classify them into 16 types based on the semantic relations. This paper includes experiments which were performed on these 16 different types. The experiments were designed following Johnson (2008); and the native speakers' intuition was measured with two scales, numerical estimates and line drawing, though the latter was adopted in the actual analyses. Through the experiment, the following facts were observed: (i) the grammaticality of the MCCs varies depending on their semantic relations, (ii) MNCs were more grammatical than MACs if both constructions occurred in similar environments, and (iii) the sentences in some MAC types had much lower grammaticality than those in the others, as Ryu (2013b, 2013c) mentioned. (**Chungnam National University & Hannam University**)

Key words: Multiple Nominative Construction, Multiple Accusative Construction, Multiple Case Construction, experimental approach, semantic relations

1. Introduction

Multiple Nominative Constructions (MNCs) and Multiple Accusative Constructions (MACs) are some of the hottest and interesting topics in Korean syntax. As Yoon (2004) pointed out, they are some of the more puzzling phenomena in topic-prominent languages such as Korean.

* Chungnam National University. College of Humanities, 99 Daehak-ro, Yuseong-gu, Daejeon, 305-764, Korea, Phone: +82-2-821-5331, E-mail: ylee@cnu.ac.kr

When the example sentences of these constructions were given to some university students in Korea, however, their grammaticality judgments were drastically different depending on which sentences were given to them. For example, let's see the following two example sentences.[1],[2]

(1) **N01**: integral object-component

Thokki-ka kwi-ka kil-ta.
rabbit.NOM ear.NOM be-long.DECL

'The ears of rabbits are long.'

(2) **N02**: collection-member

I hamtay-ka camswuham-i manh-ta.
this fleert.NOM submarine.NOM be-plenty.DECL

'There are plenty of submarines in this fleet.'

Both sentences contain MNCs, and they are examples extracted from Ryu (2013c). Each sentence has two NPs: *Thokkili-ka* and *kwi-ka* in (1) and *I hamtay-ka* and *camswuham-i* in (2). Let's call the first NP NP1 and the second NP NP2 respectively. In spite of the structural similarity of these two sentences, most of the Korean students answered that (1) was much better than (2), and more than half of them said that (2) was ungrammatical.

These examples demonstrate that native speakers' intuition is different even within the same MNCs. Then, here comes a naturally-arising question: what makes these differences in MNCs and MACs.

The goal of this paper is to investigate this question. There have been lots of previous studies on MNCs and MACs. However, Ryu (2010, 2013a, 2013b, 2013c) recently tried to unify MNCs and MACs into Multiple Case Constructions (MCCs) and to provide a unified account for them. He also classified the types of MCCs into 16 different types based on the semantic relations. In this paper, experiments were designed and performed based on these 16 relation types. The experiments were designed following Johnson (2008); and the native speakers' intuition was measured with two scales, numerical estimates and line drawing, though the latter was adopted in the actual analyses.

This paper is organized as follows. Section 2 reviews previous studies on MCCs in Korean and on the empirical/experimental methods in syntactic research. Section 3 mentions the research methods and procedure taken in this paper. Section 4 includes the analysis results, and Section 5 contains discussions and implications based on the analysis results. Section 6 summarizes and concludes this paper.

---

[1] The nominative case markers *-ka* and *-i* and the accusative case markers *-lul* and *-ul* are allomorphs, respectively. The former is post-vowel and the latter post-consonantal. The Yale Romanization System is used for the romanization of the Korean words. The abbreviations for the glosses used in this paper are as follows: NOM (nominative), ACC (accusative), DAT (dative), PRES (present tense), PAST (past tense), DECL (declarative).

[2] Here, N01 and N02 refer to the type of semantic relations. The sentence (1) has a *whole-part* relation and (2) a *class-member* relation.

2. Previous Studies

2.1 MNCs and MACs in Korean

Since Case markers are one of the typical syntactic phenomena in Korean, there have been lots of previous studies on this topic. Among the numerous previous studies, the following are included: Yang (1972), Yoon (1986), O'Grady (1991), Schütze (1996), Kim (2001), Park (2001), Park (2005), Kim, Sells, and Yang (2007), Yoon (2009), and Ryu (2010, 2013a, 2013b, 2013c). Though it is important to provide a critical review for these studies, this paper doesn't include the review in detail. Instead, this paper contains just a brief overview of previous approaches on MCCs, which is enough to proceed the discussions in Section 5.

Previous studies on Case markers are divided into roughly two groups.[3]

One is syntactic approaches and the other is semantic approaches. Syntactic approaches, once again, can be divided into two types: Constituent Approaches and Non-constituent Approaches. Constituent Approaches are based on the concept of possessor-raising or genitive NP. In this approach, NP has a structure [NP NP1 NP2] where NP1 becomes a possessor and NP2 is a possessee. Then, NP1 moves out from the NP, and the Case marker of NP1 changes into the Nominative marker *-ka* or the Accusative marker *-lul*. Many analyses including Choe (1987), Kitahara (1993), Ura (1996), and Cho (2000) took this approach.

Non-constituent Approaches have two different types of analyses. The first one is Major Subject Analyses. This approach assumes that Korean may have sentential predicates and that this language has a major subject in addition to the usual subject position. In this type of analysis, both NP1 and NP2 can be the subjects, and various notions of subjects are defined. In fact, this type of analysis started from Choe (1937), where he called them a *big subject* and a *small subject* respectively. Recently, Yoon (2003, 2009) and Lee (2007) took this approach. The second type is Topic/Focus Analyses. In these types of analyses, only NP2 is a subject and NP1 becomes a topic or a focus. Hong (1991), Rhee (1999), Yoon (1986), Schütze (2001), Kim (2000, 2001, 2004), Kim and Sells (2007), and Kim *et al.* (2007), Park (2005), Choi (2012) adopted this approach.

In contrast to syntactic approaches to MCCs, semantic approaches have focused on the licensing issues. That is, the semantic approaches to these constructions have tried to uncover what semantic relations hold between NP1 and NP2. Yang (1972) explained MCCs with *macro-micro* relations. Here, the *macro-micro* relation refers to a relation where an NP is conceptually divided into the whole NP itself and a subpart of it. The NP which corresponds to the former is referred to as a macro-NP, while that corresponding to the latter is referred to as a micro-NP. Maybe, the most frequently mentioned relation in MCCs is *inalienable possession*. That is, NP1 becomes an inalienable possessor of NP2 in these constructions. Many studies

[3] In fact, there is another type of syntactic approach to MCCs, though it has not been discussed that much in a lot of literature. That is Case Spreading Analysis in Role and Reference Grammar (RRG; van Valin and Foley, 1980; van Valin and LaPolla, 1998). In the RRG account of MCCs, Nominative/Accusative Case markers can spread from one point to the other direction. Park (1995), Han (1999), and van Valin (2009) provide this type of account to Korean MNCs and MACs. For an example sentence, see the footnote 7.

took this approach including Kang (1987), Choe (1987), Kim (1989, 1990), Yoon (1989), Maling and Kim (1992), Kitahara (1993), Yoon (1997), and Moon (2000). However, Park (2001) pointed out that not only inalienable (physical) possession but also non-physical abstract possession is also possible in MCCs, and he proposed *g(eneralized)-possession* in order to cover both kinds of possession relations.

On the other hand, Na & Huck (1993) proposed (*thematically*) *subordinate condition*. Accordingly to them, X is *thematically subordinate* to an entity Y iff Y's having the properties that it does entail that X has the properties that it does. Kim (2000, 2001, 2004), Kim and Sells (2007), and Kim *et al.* (2007) adopted this condition and tried to capture the semantic relations between NP1 and NP2.

Several other studies tried to capture the relation with *aboutness* condition. This condition can be defined as follows: if an element is characterized by the subsequent phrase, it satisfies the *aboutness* condition. Kang (1988), O'Grady (1991), Hong (1997), Yoon(2004), Choi and Lee (2008), and recently Choi (2012) adopted this approach. Recently, Ryu (2013a, 2013b, 2013c) recently tried to unify MNCs and MACs into Multiple Case Constructions (MCCs) and to provide a unified account for them. He also classified the types of MCCs into 16 different types based on the *conceptual linking hierarchy*.

Even though there are a lot of studies on the theoretical accounts for MCCs, only a few provided the classifications of MCCs in Korean, such as Yang (1972), Na and Huck (1993), and Park (2001). Recently, Ryu (2013b: 192) summarized these classifications as follows.

Proposed type of MCCs	NOM-NOM	ACC-ACC	Yang (1972)	Na & Huck (1993)
Type 01 integral obj.-component	◯	◯	whole-part	meronomic rel.
Type 02 collection-member	◯	◯	×	×
Type 03 mass-portion	◯	◯	×	×
Type 04 object-stuff	◯	◯	×	×
Type 05 activity-feature	◯	◯	×	×
Type 06 area-place	◯	◯	×	×
Type 07 class-membership	◯	◯	class-member type-token	taxonomic rel.
Type 08 object-attachment	◯	◯	×	×
Type 09 object-quality	◯	◯	×	qualitative
Type 10 object-quantity	◯	◯	total-quantity	×
Type 11 space-object	◯	*	×	×
Type 12 time-object	◯	*	×	×
Type 13 possessor-object	◯	*	×	
Type 14 conventional relation	◯	*	×	conventional
Type 15 object-predication	◯	*	×	×
Type 16 conversive relation	◯	*	affected-affector	conversive

[Table 1] Types of Multiple Case Marking Constructions (Ryu (2013b: 192))

The first column enumerates the types of semantic relations, which holds between NP1 and NP2. Ryu (2013a, 2013b, 2013c) re-organized the classifications based on previous studies such as Yang (1972), Na and Huck (1993), and Park

(2001). Some of the type names come from the previous studies, and others were made by him. The second and third column demonstrates if these types occur in the MNCs and MACs. Here, the symbol ◯ refers to 'possible' and * to 'impossible'. The last two columns show us how each semantic relations were referred to in Yang (1972) and Na & Huck (1993) respectively. Here, the symbol × refers to 'not mentioned'. And, rel. and con. are abbreviations of *relation* and *constructions* respectively.

The criteria of these classifications are the semantic relations which hold between the two consequative NPs, i.e. the semantic relations between NP1 and NP2. He also provided example sentences for these types.

2.2 Grammaticality Judgment Task

A grammaticality judgment task (also known as native speakers' intuition test) is a psychological experiment which can be used to get the subconscious knowledge of native speakers in a given language. It involves asking native speakers to read a sentence and judge if it is well-formed (grammatical), marginally well-formed, or ill-formed (unacceptable or ungrammatical) (Carnie, 2012).

As Johnson (2008:218) mentioned, in syntactic research, an interval scale of grammaticality is commonly used. There are usually five steps of scales, and sentences are rated by native speakers as grammatical (no mark), questionable (? or ??), and ungrammatical (* or **). This is essentially five-point category rating scale, and the researcher could give people this rating scale and average the test results, where **=5, *=4, ??=3, ?=2, and no mark=1. However, it has been observed in the study of sensory impressions that raters are more consistent with an open-ended ratio scale than they are with category rating scale (Stevenson, 1975). Recently, researchers have had an interest in native speakers' intuition on syntactic data (Bard, Robertson, and Sorace, 1996; Schütze, 1996; Cowart, 1997; Keller, 2000). So, in recent years, various methods have been adapted into the study of sentence acceptability, from the study of psychophysics which studies the subjective impressions of physical properties of stimuli.

Johnson (2008) adopted a technique, so called *magnitude estimation*, using an open-ended ratio scale for reporting the impressions of native speakers. The experiment in his proposal starts with a demonstration of magnitude estimation by asking participants to judge the length of a few lines. These practice judgments provide a sanity check in which we can evaluate the participants' ability to use magnitude estimation to report their impressions. Stevenson (1975) found that numerical estimates of line length have a one-to-one relationship with actual line length (that is, the slope of the function relating them is 1). In the second session, the participants were presented with sample sentences. Some are grammatical and the others are not. Then the participants were instructed to judge how good or bad each sentence is by drawing a line that has a length proportional to the grammaticality of the sentence. In the third session, the participants were provided the target sentences. Their job was to estimate the grammaticality of the target sentences by drawing lines, which indicate native speakers' impression of the grammaticality with the length of the line which they draw for the sentences. In the last session, the participants were provided the same target sentences. They were asked to estimate

the grammaticality of the target sentences with numerical estimations, which also indicate their impression on the acceptability of the target sentences.

Lodge (1981) mentioned that this *magnitude estimation* has three advantages over category scaling. First, the latter has limited resolution. For example, if native speakers may feel that a sentence is somewhere between 4 and 5 (something like 4.5), gradient ratings are not available in the latter method. However, the former permits as much resolution as the raters wish to employ. Second, the latter method uses an ordinal scale, and there is no guarantee that the interval between * and ** represent the same difference of impressions as that between ? and ??. The former method, on the other hand, provides judgments on an interval scale for which averages (mean value, m) and standard deviations (s) can be more legitimately used. Third, the latter limits our ability to compare results across the experiments. The range of acceptability for a set of sentences has to be fitted to the scale, and what counts as ?? for one set of sentences may be quite different from what counts as ?? for another set of sentences.

However, *magnitude estimation* also has some shortcomings. In Johnson's proposal, for example, the participants in the example experiment were asked to judge sentences into two ways: (1) by giving a numeric estimate of acceptability for each phrase, as they did for the lengths of lines in the practice session; and (2) by drawing lines to represent the acceptability of each line. Bard *et al.* (1996) found that the participants sometimes think of numeric estimates as something like academic test scores, and so they limit their responses to a somewhat categorical scale (e.g. 70, 80, 90, 100), rather than using a ratio scale as intended in the magnitude estimation. Consequently, the participants have no such preconceptions about using a line length to report their impressions, and we might expect more gradient unbounded responses by measuring the lengths of lines that participants draw to indicate their impressions of sentence grammaticality.

3. Research Methods

3.1 Experimental Design

There are usually two types of strategy to get the linguistic data in syntactic literature. One is corpus and the other is a grammaticality judgment task.

The first option is to use corpus data. A corpus is a collection of spoken or written texts. There are also some corpora in Korean. However, there are some problems with using the corpus data in this study. First, there is no special corpus in which the data were collected only for Case phenomena in Korean. Accordingly, it is impossible to find a specialized corpus containing the data only for MCCs. Second, even if a general-purpose corpus is available, there are some problems with using it in this study. Among the corpora authentic and widely used in the studies in Korean, there is the *Sejong Corpus* (the Korean National Corpus; Seo, 2002). However, this corpus contains very few sentences showing MCCs. Therefore, it is questionable to use the data in quantitative and statistical approaches. There is another problem. Even though the corpus data contain the sentences showing MCCs, not all the variations occur in the corpus. Accordingly, it is unreasonable to use the corpus data in this study.

The second option is to perform a grammaticality judgment task. As mentioned in Section 2.2, it is a psychological experiment which can be used to get the subconscious knowledge of native speakers in the given language. This paper basically adopted this option and followed the experimental design described in Johnson (2008). The experiments were conducted two times in the fall semester in 2013 for the purpose of consistence testing with the same participants. One is for MNCs and the other is for MACs. The participants were registered in university (freshmen and sophomores) at the time of the experiments, who were not linguistics majors. A total of 27 students participated in the experiments.

Each experiment was performed as follows. Each experiment consisted of four sections, following Johnson (2008). In the first section, the participants were given a sample line, and the numerical score of 100 was given to the line. Then, they were provided with 10 lines with different length, and they were instructed to judge the length of the lines. They were said to write the numerical estimates for each line, which they thought of as the lengths of the lines compared with the standard line with the numerical score of 100. In the second section, they were given a sample Korean sentence perfectly grammatical. The numerical estimate 157 was given to the sentence. This value was given to the participants in order to avoid the same problem that Bard *et al.* (1996) pointed out. Then, they were provided with 10 different Korean sentences. Some of them were grammatical, some others were ungrammatical, and the others are in-between. They were instructed to draw a line for each sentence which corresponded to their judgment on the acceptability, compared with that of the standard line. The possible length of the lines ranged from 0 mm to 170 mm.[4]

In the third section, the target sentences were given. The participants were instructed to estimate the grammaticality of the target sentences by drawing lines. The possible length of the lines ranged from 0 to 170 mm, as in the second section. In the last session, the participants were provided with the same target sentences. Now, they were to estimate the grammaticality of the target sentences with numerical estimations. The possible range of numerical scores was from 0 to 200.

This paper used the target sentences which were contained in Ryu (2013c). There are two reasons to use the sentences in Ryu (2013c). First, Ryu (2013c) contained almost all of the MCC types, which were mentioned in previous researches include Yang (1972), Na and Huck (1993), and Park (2001). Therefore, it was possible to have a bird's eye view to MCCs in Korean. Second, Ryu (2013c) provided the sentences which belonged to both MNCs and MACs. Accordingly, it was easy to get the target sentences for both constructions. The experiments in this paper used the target sentences without any modification in order to avoid any irrelevant distortion when the lexical items were changed.

Example sentences for MNCs were given in (1) and (2). Likewise, the target sentences for MACs were also extracted from Ryu (2013c). The following sentences

[4] The possible lengths of the lines ranged from 0 to 190 mm if the participants used them up to the right margin. However, the maximum value was 169 mm in all of the experiments.

in (3) and (4) are the counterparts of (1) and (2) (Ryu, 2013c:11).[5].[6]

(3) **N01**: integral object-component

Hans-ka thokki-lul kwi-lul cap-ass-ta.
Hans.NOM rabbit.ACC ear.ACC grab.PAST.DECL

'Hans grabbed the ears of rabbits.'

(4) **N02**: collection-member

Cekkwun-i i hamtay-lul camswuham-ul paksalnay-ss-ta.
enemy.NOM this fleert.ACC submarine.ACC destroy.PAST.DECL

'The enemy destroyed the submarines of this fleet.'

Since both MNCs and MACs had 16 types, a total of 32 target sentences were used in the experiment. Along with these target sentences, distracting sentences of the same number (16 sentences) were also provided for MNCs and MACs respectively, unlike Johnson (2008). Accordingly, a total of 64 sentences were included in the experiment. Then, the collected sentences were randomly ordered and provided to the participants.

After the experiments, the 32 sets of data were extracted from the target sentences: 16 for MNCs and 16 for MACs. For each experiment, students of different numbers participated, since the experiments were performed two times. Accordingly, the data sets were extracted only from those students who participated in all of the experiments. Among the 27 students, only 23 participants participated in both experiments. However, among the answers of these 23 students, some answers were missing. That is, there were some students who answered to some sentences but provided no answer to some others. Two students answered in this fashion, and the data sets for these students were excluded. Finally, the data sets of the remaining 21 participants were extracted. However, among those students, one

[5] Though Ryu (2013c) mentioned that the sentences in **Type 11-Type 16** in MACs were ungrammatical, the experiment in this paper includes the sentences in order to check how much their grammaticality was bad to native speakers.

[6] It doesn't imply that MACs have parallel structures with MNCs and that they have to be analyzed with the same mechanism with MNCs. Let's see the following sentences (Han, 1999).

(i) a. Chelswu-ka Yenghi-eykey kkoch-ul cwu-ess-ta.
Chelswu.NOM Yenghi.DAT flower.ACC give.PAST.DECL
'Chelswu gave a flower to Yenghi.'

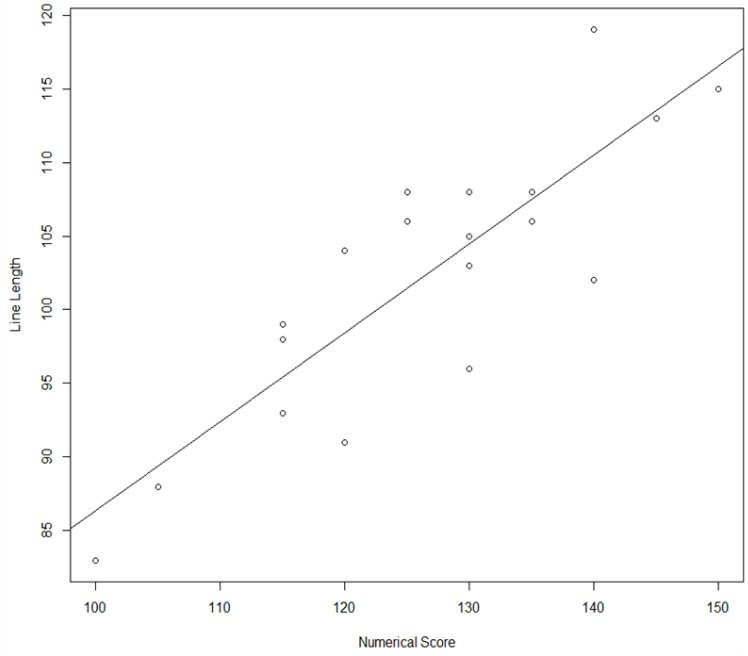
b. Chelswu-ka Yenghi-lul kkoch-ul cwu-ess-ta.
Chelswu.NOM Yenghi.ACC flower.ACC give.PAST.DECL
'Chelswu gave Yenghi a flower.'

In the RRG account, this sentence can be explained with Case Spreading. That is, the Accusative marker *-lul* spreads to the left, and the Dative Case marker *-eykey* in (ia) is changed into an Accusative in (ib). Though both (3)/(4) and (ib) contain MACs, their sources are different. In our terminology, the two NPs have different semantic relations in (3) and (4). It may be impossible to improvise the MNC counterpart of the sentence in (ib). As this sentence illustrates, MACs may have different syntactic structures and semantic relations from MNCs.

student belonged to the outlier in terms of age. Accordingly, the data sets for the student were also excluded.[7]

Consequently, the data sets for only a total of 20 students were included in the statistical analyses.[8] The age distribution of those 20 students was as follows: m=20.15, s=0.93. Though the majority of the students were 19 and 20, some male students were over 21.[9]

For each student, 32 target sentences were collected (16 for the MNCs and 16 for MACs). For each of the data sets, two different kinds of data were extracted: one for *numerical estimates* and the other for *line drawing*. Since we had two different kinds of scales, it was necessary to check the correlation between these two scores. Figure 1 shows the correlations of the first data set, which was shown in (1).



[Figure 1] Correlation between Numerical Score and Line Length

Here, r was 0.853. Since it is said that the two variables have correlations if the r value is over 0.5, it will be safe to say that the *line drawing* and *numerical estimate* are highly correlated in this data set.

In the actual statistical analyses below, the scores for the *line drawing* were used. The reason was that the problem of category scaling can be avoided in the scores for line drawing. Even though the participants were given the 0-200 numerical ranges, they used only some of them, i.e., the multiple numbers of 5 or 10. In

[7] The data sets for the student were excluded also for easy calculation of percentage in Section 5.

[8] All the statistical analyses in this paper were performed using R.

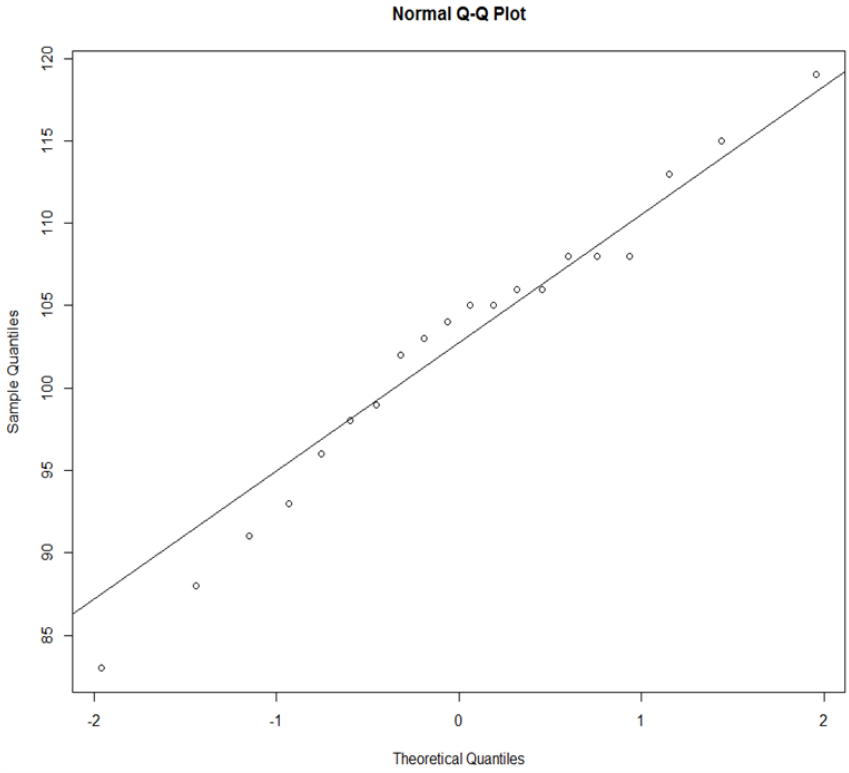
[9] In Korea, male students usually serve their military service for almost two years. That's why some male students are older than female students.

the scores for the line drawing, since they were instructed to draw a line without a ruler, they would avoid such kind of subconscious tendency. Because the line lengths were highly correlated with the numerical estimates, it was possible to use only the scores for the line drawing in the analyses.

3.2 Normality Test

After we got the scores for the *line drawing* for each target sentence, the first thing that we had to do was a normality test. The reason was that the types of the following statistical test were determined by the results of the normality tests. If the distributions of data followed the normal distribution, we could apply parametric tests such as a *t-test* or an ANOVA. If not, non-parametric tests had to be applied including Wilcoxon tests or Friedman tests. Therefore, it was important to check if the distributions of data sets followed the normal distribution or not.

There are a few different sorts of normality tests. One is to use a Normal Quantile Plot (Baayen, 2008). For example, the 20 data for sentence (1) can be represented in the Normal Quantile Plot as follows:



[Figure 2] Normal Quantile Plot for Sentence (1)

In this plot, the closer the points get to the Q-Q line, the closer they are to the normal distribution. As you can see, most of the points, especially those in the middle, are attached very close to the Q-Q line, Accordingly, we can guess that this data follows the normal distribution.

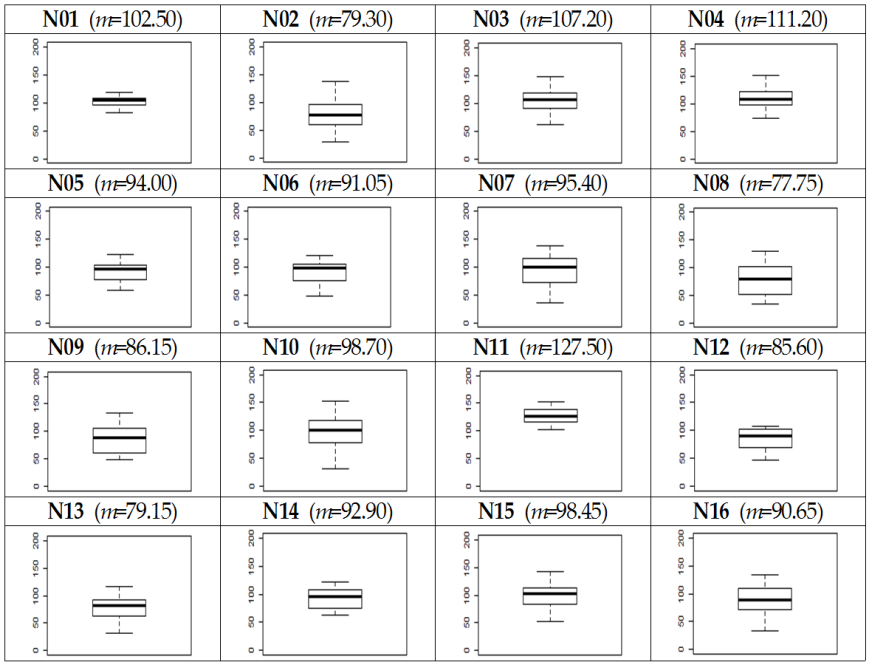
However, one of the disadvantages using the Normal Quantile Plot is that we cannot numerically decide if the given data follows the normal distribution or not. The normality test that solves this problem is a Shapiro-Wilk Normality Test. For example, if we perform the test with the scores for sentence (1), we have a p-value 0.825. Since this p-value is much bigger than the a-value of 0.05, we cannot reject the Null Hypothesis that this data follows the normal distribution. That is, we can say that this data follows the normal distribution.

In the actual statistical analyses, Shapiro-Wilk Normality Tests were used. If the p-value is bigger than the a-value of 0.05, the data is said to follow the normal distribution. If the p-value is smaller than the a-value of 0.05, the data is said not to follow the normal distribution. In our data, only one set of data (**A04** in Table 4) didn't follow the normal distribution. Accordingly, parametric tests were frequently used such as t-tests or ANOVA. However, non-parametric tests were also applied when they were necessary, including Wilcoxon tests or Friedman tests.

4. Analysis Results

4.1 MNCs

Table 2 illustrates the results of the grammaticality judgment task for the 16 types of MNCs. For each type, the mean values are provided in addition to the box plots.



[Table 2] Results of the Grammaticality Judgment Task for the 16 Types of MNCs

In order to examine if the mean values became different depending on the semantic relations, a statistical test had to be performed. For the test, the answers for each participant were paired with another. Let's say that P01.N01 refers to the

score for the **N01** of the first participant (P01). For example, P09.N07 refers to the score for the **N07** of the 9th participant (P09). Now, P01.N01 makes a group with P01.N02, P01.N03, ... , P01.N16. All the other 19 groups were made with the same mechanism. The purpose of this testing was to examine if each native speaker would judge the grammaticality of sentences differently depending on the semantic relations. Since all the types in MNCs follow the normal distribution, a repeated-measures ANOVA was performed, and the result was that the mean values became significantly differentiated depending on which semantic relations MNCs had (F=6.818, p<0.001).

Next, in order to examine the mean value of which type was significantly different from that of which type, a Tukey's HSD test (the parametric post-hoc test) was performed, and its results are shown in Table 3. Here, '×' is used when 0.05<p, '*' when p<0.05, '**' when p<0.01, and '***' when p<0.001.

	N01	**N02**	**N03**	**N04**	**N05**	**N06**	**N07**	**N08**	**N09**	**N10**	**N11**	**N12**	**N13**	**N14**	**N15**
N02	×														
N03	×	**													
N04	×	***	×												
N05	×	×	×	×											
N06	×	×	×	×	×										
N07	×	×	×	×	×	×									
N08	*	×	**	***	×	×	×								
N09	×	×	×	*	×	×	×	×							
N10	×	×	×	×	×	×	×	×	×						
N11	*	***	×	×	***	***	***	***	***	**					
N12	×	×	×	×	×	×	×	×	×	×	***				
N13	×	×	**	***	×	×	×	×	×	×	***	×			
N14	×	×	×	×	×	×	×	×	×	×	***	×	×		
N15	×	×	×	×	×	×	×	×	×	×	×	×	×	×	
N16	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

[Table 3] Results of the Tukey's HSD Test for the 16 Types of MNCs

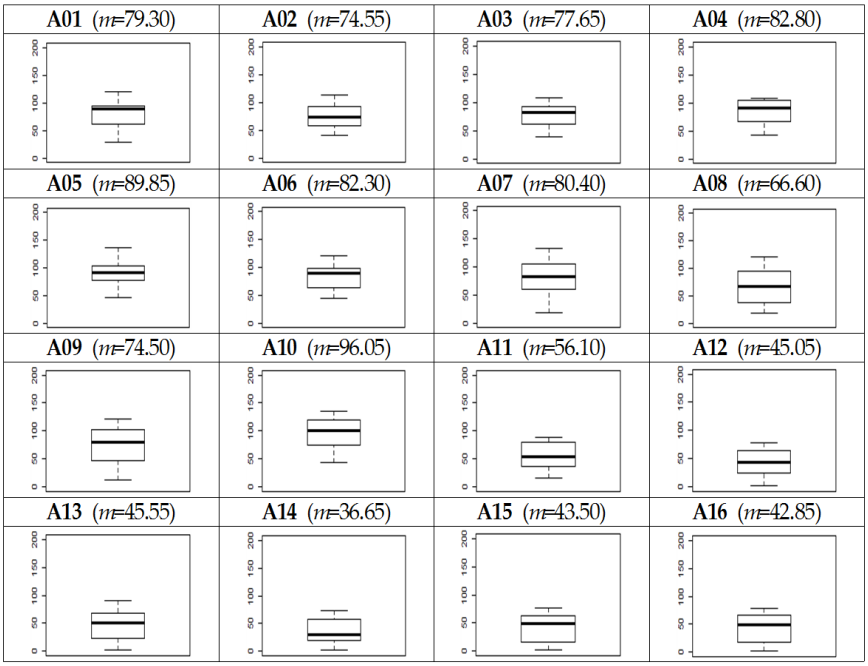
Among the 120 pairs (=16×(16-1)/2), 19 pairs (15.83%) had statistically significant differences.

Here, remember that the following observation was mentioned in Section 1: most Korean students answered that (1) was much better than (2), and more than half of the students said that (2) is ungrammatical. This tendency can be predictable from the table. Note that the mean value of **N01** was 102.50 and that of **N02** was 79.30. However, the p-value of the **N01**-**N02** pair was 0.083. Since this value is bigger than 0.05, we cannot say that the two groups had significantly different mean values. However, since the p-value was too close to the a-value, there is still a possibility that the differences between (1) and (2) were not made by chance but by a systematic factor.[10]

[10] Note that the Null Hypothesis would be rejected if the a-value were 0.1 (not 0.05). Then, we have to say that the two groups have significantly different mean values. That is, the differences between (1) and (2) were not made by a systematic factor, not by chance.

4.2 MACs

Table 4 illustrates the results of the grammaticality judgment task for the 16 types of MACs. As in MNCs, the mean value is provided in addition to the box plot for each type.



[Table 4] Results of the Grammaticality Judgment Task for the 16 Types of MACs

If each type in MACs was compared with that of MNCs, the following two facts were observed. First, the overall mean values of MACs were lower than those of MNCs. Second, the minimum scores of some types are very close to 0, especially from **A12** to **A16**.

In order to examine if the mean values became different depending on the semantic relations, the data sets were made similarly as in the MACs. Even though one type didn't follow the normal distribution (**A04**), a repeated-measures ANOVA was performed and the result was that the mean values became significantly differentiated depending on what semantic relations MNCs had (F=11.01, p<0.001).[11]

Next, in order to examine the mean value of which type was significantly different from that of which type, a Tukey's HSD test was performed, and its results are shown in Table 5.

Among the 120 pairs, 45 pairs (37.50%) had the significant differences. However, note that the significances occurred in the lower parts of the table, mainly

[11] Strictly speaking, a Friedman test (the non-parametric counterpart of the repeated-measures ANOVA) had to be used here. However, since only one (**A04**) didn't follow the normal distribution among 16 data sets, a repeated-measures ANOVA was used here. In fact, there was no differences in the analysis results (χ^2=185.97, p<0.001). However, if 3 or 4 data sets didn't follow the normal distribution, a Fiedman test has to be used.

	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**	**A13**	**A14**	**A15**
A02	×														
A03	×	×													
A04	×	×	×												
A05	×	×	×	×											
A06	×	×	×	×	×										
A07	×	×	×	×	×	×									
A08	×	×	×	×	×	×	×								
A09	×	×	×	*	×	×	×	×							
A10	×	×	×	×	×	×	×	*	×						
A11	×	×	×	×	**	×	×	×	×	***					
A12	**	×	**	***	***	**	**	×	*	***	×				
A13	**	×	×	**	***	**	**	×	*	***	×	×			
A14	***	***	***	***	***	***	***	*	***	***	×	×	×		
A15	**	*	**	***	***	***	**	×	*	***	×	×	×	×	
A16	**	*	*	***	***	*	***	×	×	***	×	×	×	×	×

[Table 5] Results of the Tukey's HSD Test for the 16 Types of MACs

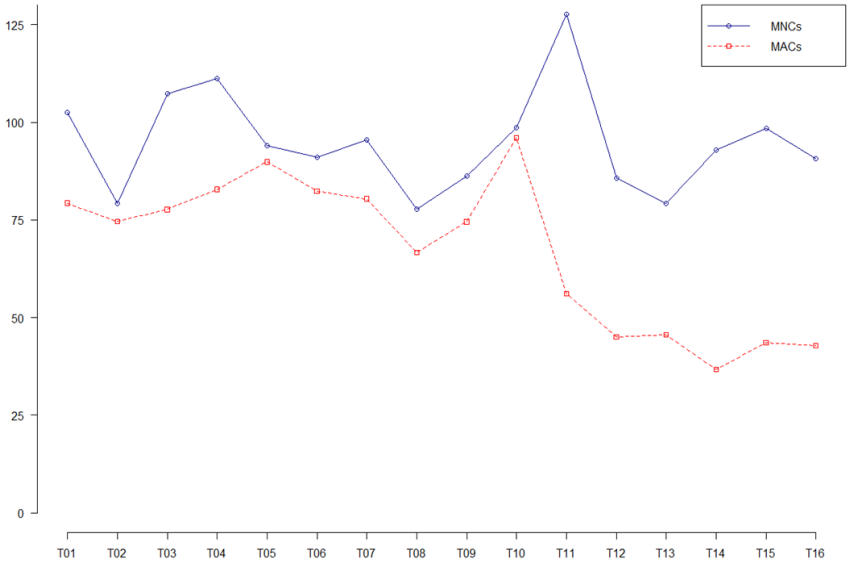
from **A11** to **A16**. This means that the grammaticality distribution can be divided roughly into two groups. One is from **A01** to **A10**, and the other is from **A11** to **A16**.

In order to examine if there are statistically significant differences between the two groups, the data were collected into two groups separately. Group 1 was composed of the data from **A01** to **A10**, and Group 2 was from **A11** to **A16**. When Shapiro-Wilk Normality Tests were performed for these two groups of data, both groups didn't follow the normal distributions (Group 1: $p<0.001$; Group 2: $p<0.001$). Since both groups didn't follow the normal distributions, a Mann-Whitney's U test (the non-parametric counterpart of an independent sample t-test) was performed, which was also known as a Wilcoxon Rank Sum Test. The results was that there are statistically significant differences between the two groups ($W=19796.5$, $p<0.001$). This results indirectly supports Ryu's claim (2013b, 2013c) that the sentences from **A11** to **A16** are different those in other groups.

4.3 MNCs vs. MACs

Now, let's see how the semantic relations affected the grammaticality of MNCs and MACs. Figure 3 shows the comparison of the scores in MNCs and MACs.

As you can observe, there are some differences between each pair of types. In order to examine if the distributions of MNCs were different from those of MACs, each answer in MNCs was paired with that of MACs. That is, P01.N01 was paired with P01.A01, P01.N02 with P01.A02, and so on. Since both data sets followed the normal distributions except **A04**, paired t-tests were performed. The results showed that the distributions of MNCs were significantly higher than those of MACs ($t=11.99$, $p<0.001$). The analysis results for each pair are shown in Table



[Figure 3] MNCs vs. MACs

7.[12]

	T01	**T02**	**T03**	**T04**	**T05**	**T06**	**T07**	**T08**
t	3.689	0.616	4.383	4.357	0.691	1.247	1.504	1.207
p	0.001	0.542	**0.000**	**0.000**	0.494	0.220	0.141	0.235
	T09	**T10**	**T11**	**T12**	**T13**	**T14**	**T15**	**T16**
t	1.348	0.298	10.97	6.063	4.236	8.606	7.106	6.652
p	0.186	0.767	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

[Table 6] MNCs vs. MACs

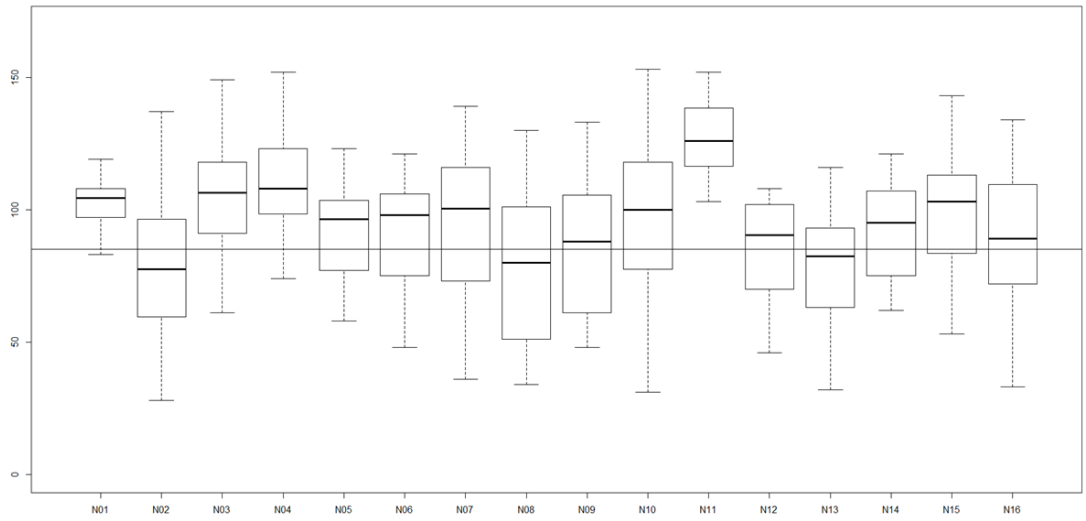
Here, the p-value is bold-faced when it is less than 0.05, which means significant differences. As you can observe **N01** is different from **A01**, **N03** from **A03**, **N04** from **A04**. Also note that and **N11-N16** are different from **A11-A16**. As evidenced in all these data, we can say that the distributions of half of MNCs are significantly different from those of MACs.

5. Discussions

In this paper, it was examined how native speakers' intuition varies depending on the semantic relations of MCCs. From these experiments, we may ask the following question: Does the Korean language really have MNCs and MACs? To answer this question, let's think about the analysis results in the experiment. In this experiment, the effects of semantic relations were investigated.

[12] Likewise, a Wilcoxon test (the non-parametric counterpart of the paired t-test) had to be used for the N04-A04 pair. However, there was no difference in the analysis result (V=173, p=0.002).

Let's see the MNCs first. As mentioned in Section 4.1, all of MNCs' examples in the experiment came from Ryu (2013c). These data contained the typical MCC relation types which were frequently mentioned in previous studies such as Yang (1972), Na and Huck (1993), and Park (2001). Notwithstanding, the grammaticality judgments on these typical examples in MNCs were not identical. Figure 4 demonstrates the grammaticality judgments of MNCs in the experiment.



[Figure 4] Grammaticality Judgments of MNCs depending on the Semantic Relations

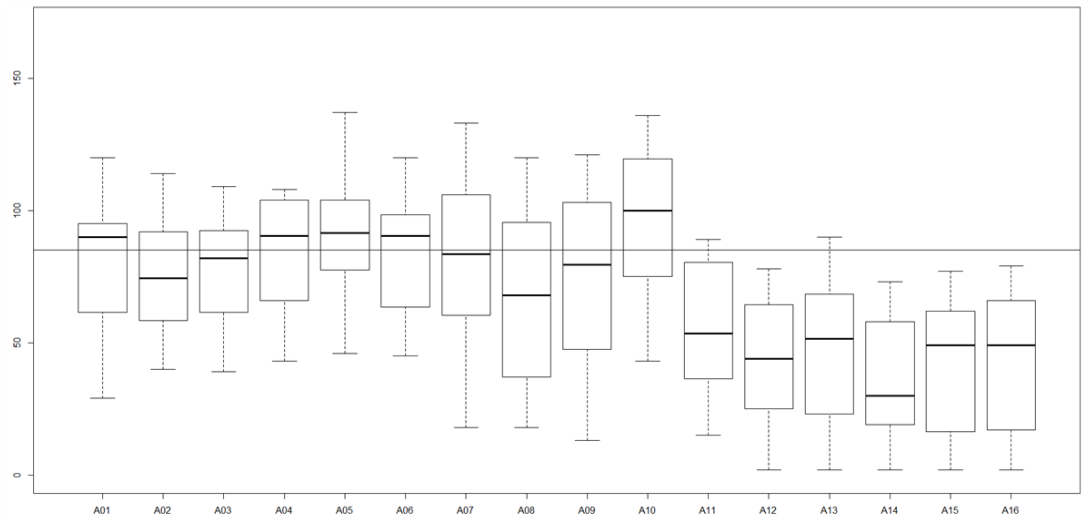
Here, the line in the middle is the one for the line length 85. As mentioned in Section 3.1, the maximum length of lines allowed for the participants to draw was 170 mm. Therefore, the line for the line length 85 becomes the baseline which divides the grammaticality judgments into just two parts: toward grammatical part and toward ungrammatical part. Let's call them the positive zone (85 mm - 170 mm) and the negative zone (0 mm - 84 mm) respectively.

As you can see, the mean values of only the three types (**N02**, **N08**, and **N13**) are located in the negative zone, while the values for the other types are in the positive zone. And, in only two types (**N06** and **N11**), even the minimum values are in the positive zone. These facts mean that MNCs surely exist in Korean.

Then, how about the MACs? Figure 5 demonstrates the grammaticality judgments of MACs in the experiment.

As you can see, the mean values of most types are located in the negative zone. Even though the sentences from **A11** to **A16** are excluded from the discussion, the mean values of five types (**A01**, **A04**, **A05**, **A06**, and **A10**) are located in the positive zone. This implies that most of native speakers are not sure of the grammaticality of MACs.

Then, the next question is what the analysis results imply in the experiments. The box plots in Figure 4 and Figure 5 have the following implications in the studies of MCCs.



[Figure 5] Grammaticality Judgments of MACs depending on the Semantic Relations

First, the distributions of data in Figure 4 and Figure 5 indicate that detailed examinations of syntactic phenomena have to come before their theoretical explanations of MCCs. Most previous approaches to these constructions have been theoretical in nature and have focused on how these constructions were made or what the semantic relations licensed their constructions. However, as you can see in Figure 4 and Figure 5, the native speakers' grammaticality was different depending on which type of sentences they were provided with. For example, even though most previous studies implicitly assumed that both (1) and (2) are grammatical, the experimental results demonstrated that the grammaticality of (1) is much higher than that of (2). In addition, the ranges/variances of the grammaticality were various from type to type. For example, N01 and N11 have small variances in MNCs, whereas N02, N07, N10, and N16 have large variances. Likewise, A03 and A04 have small variances in MACs, whereas A07, A08, A09, and A16 have large variances. This means that some extent of agreements can be drawn from the former groups but that those kinds of agreements cannot be drawn from the latter groups.[13] Accordingly, more studies are necessary on which factors would make these discrepancies in the native speakers' grammaticality to MCCs.

Second, the data show that the example sentences had to be used with a caution when they were given to the students in Korean. For example, if the example sentences belonged to the types of N01 and N11 in MNCs, most students would say that the sentences were grammatical. However, if the example sentences were within the types of N02 or N08, most of people would say that the sentences

[13] This is why the data were first provided with the box plots in Section 4.1 and Section 4.2. However, there were some cases where bar plots or line plots were more comfortable for comparison. This is why the line plot was also used in Section 4.3.

were ungrammatical. The tendencies are also applicable to MACs. if the example sentences belonged to the types of A05 and A10 in MACs, most students would say that the sentences were grammatical. However, if the example sentences were within the types of A02 or N08, most of people would say that the sentences were ungrammatical. Therefore, the answer of the students would be different depending on which types of sentences were provided to them.

Third, the data in Figure 4 and Figure 5 imply that the semantic approaches are more appropriate than the syntactic approaches to provide accounts for MCCs. As mentioned in Section 2.1, previous approaches to these constructions can be divided into two types: syntactic approaches and semantic approaches. Syntactic approaches have focused on how these constructions are constructed and they are divided into two groups (Constituent Approaches and Non-constituent Approaches). Semantic approaches to these constructions have focused on semantic relations between NP1 and NP2. They proposed the relations such as *macro-micro* relation, *inalienable possession*, *g(eneralized)-possession*, *subordinate* condition, *aboutness* condition, and *conceptual linking hierarchy*. Let's see the data in Figure 4 again. As you can see, some types are more acceptable than some others. In addition, the ranges/variances of the acceptability were various from type to type. N01 and N11 have small variances, whereas N02, N07, N10, and N16 have large variances. Likewise, A03 and A04 have small variances in MACs, whereas A07, A08, A09, and A16 have large variances. How can the syntactic approaches explain these discrepancies between the sentences, whether it is a Constituent Approach or a Non-constituent Approach?[14] However, the semantic approaches could work. In the semantic approaches, each sentence type may have different semantic relations, and these different semantic relations decide the grammaticality of the sentences and their ranges/variances. Consequently, it can be said that semantic approaches are more appropriate than syntactic approaches to provide accounts for MCCs.

Fourth, the box plots in Figure 4 and Figure 5 imply that *magnitude estimation* has advantages over *category scaling* in the grammaticality judgment tasks. For example, as you can observe in Figure 4, the range values of the grammaticality were various from type to type. N10 had the maximum range 122, and N01 had the minimum range 36. The range value of N10 is almost 4 times as big as that of N01. The differences in the range values were able to be noticed here since *magnitude estimation* was adopted in the experiments. If *category scaling* with 5 or 7 steps had been used instead, these range differences could not be observed or the differences could be smaller than the values gauged with *magnitude estimation*. This fact demonstrates that *magnitude estimation* has advantages over *category scaling* in the grammaticality judgment tasks.

6. Conclusion

In this paper, we took empirical approaches and examined how the grammaticality of MNCs and MACs in Korean varies depending on the semantic relations which hold between NP1 and NP2. We found that there were some discrepancies

[14] Note that the Case Spreading accounts in RRG cannot explain these discrepancies, either.

in the answers towards the target sentences, and then we started from the following questions: why do native speakers in Korean show different answers to these constructions. In order to solve the puzzle, three experiments were designed and performed. In the experiments, 20 university students participated. The grammaticality judgment tasks were designed following the guidelines in Johnson (2008), and native speakers' intuition was measured with two scales: *numeric estimates* and *line drawing*. After the intuition tests, the normality tests were performed on each part of the collected data. Depending on whether the data sets followed the normal distribution or not, parametric or non-parametric tests were performed.

The analysis results were as follows. First, it was found that semantic relations between NP1 and NP2 really affected the grammaticality of the sentences. Second, the grammaticality of MACs was significantly lower than that of MNCs. Third, the grammaticality of A11-A16 were significantly lower than that of A01-A10.

These analysis results have some implications that the examples should have been given after careful investigations in these constructions, because not all the sentences got positive answers from the native speakers. The results also showed that there must be systematic and scientific studies on the factors to decide the grammaticality of these sentences.

It cannot be said that the test results in this paper surely represent all the native speakers' intuition toward MCCs. However, the 20 participants satisfy the minimum requirement of the experiment. Though the number was small, it is enough to demonstrate the general tendencies in MCCs. Of course, more studies have to be performed on the data with more factors and their interactions. However, the test results in this paper surely showed the native speakers' tendency toward MCCs in Korean. Though more studies are necessary, the experimental design and the analysis methods are pre-requisite for the theoretical studies of MCCs.

References

Baayen, Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bard, Ellen, Dan Robertson, and Antonella Sorace, 1996. Magnitude estimation of linguistic acceptability. *Language* 72, 32-68.

Carnie, Andrew. 2012. *Syntax: A Generative Introduction*. 3rd Edition. Oxford: Blackwell.

Cho, Seongeun. 2000. *Three Forms of Case Agreement in Korean*. Ph.D. thesis, State University of New York, Stony Brook, New York.

Choe, Hyun-Bae 1937. *Wulimalpon* (*The Korean Grammar*). Seoul: Top Publishing Company.

Choe, Hyun-Sook 1987. Syntactic adjunction, A-chains, and the ECP: Multiple identical case construction in Korean. In Joyce McDonough and Bernadette Plunkett (eds.), *Proceedings of the 17th North East Linguistic Society* (NELS-17), 100-121.

Choi, Incheol. 2012. Sentential specifiers in the Korean clause structure. In Stefan

Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar* (HPSG-19), 75-85.

Choi, Incheol. and Eunsuk Lee. 2008. The partial parallelism between double nominative construction and ECM construction in Korean. *Studies in Modern Grammar* 53, 75-101.

Cowart, Wayne. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousands Oaks, CA: Sage Publications.

Han, Jeonghan. 1999. *On Grammatical Coding of Information Structure in Korean: A Role and Reference Grammar Account*. Ph.D. thesis. The State University of New York at Buffalo.

Hong, Ki-Sun. 1991. *Argument Selection and Case-Marking in Korean*. Ph.D. thesis. Stanford University.

Hong, Ki-Sun. 1997. Eynge-wa kwuke-uy insang kwumwun pikyo pwunsek (Subject-to-object raising constructions in English and Korean). *Language Research* 33, 409-434.

Johnson, Keith. 2008. *Quantitative Methods in Linguistics*. Oxford: Blackwell.

Kang, Beom-Mo. 1988. *Functional Inheritance, Anaphora and Semantic Interpretation*. Ph.D. thesis. Brown University.

Kang, Myung-Yoon. 1987. 'Possessor raising' in Korean. In Susumo Kuno *et al.* (eds.), *Harvard Studies in Korean Linguistics* 2, 80-88.

Keller, Frank. 2000. *Gradient in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. Ph.D. thesis. University of Edinburgh.

Kim, Jong-Bok. 2000. A constraint-based and head-driven approach to multiple nominative constructions. In Dan Flickinger and Andreas Kathol (eds.), *Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar* (HPSG-7), 166-181.

Kim, Jong-Bok. 2001. A constraint-based approach to some multiple nominative constructions in Korean. In Akira Ikeya and Masahito Kawamori (eds.), *Proceedings of the 14th Pacific Asia Conference on Language, Information, and Computation* (PACLIC-14), 165-176.

Kim, Jong-Bok. 2004. *Hankwuke Kwukwuco Mwunpep* (written in Korean, *Korean Phrase Structure Grammar*). Seoul: HankukMunhwasa.

Kim, Jong-Bok and Peter Sells. 2007. Two types of multiple nominative construction: A constructional approach. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar* (HPSG-14), 364-372.

Kim, Jong-Bok, Peter Sells, and Jaehyung Yang. 2007. Parsing two types of multiple nominative constructions: A constructional approach. *Language and Information* 11, 25-37.

Kim, Young-Joo. 1989. Inalienable possession as a semantic relationship underlying predication: The Case of Multiple-Accusative constructions. In Susumo Kuno *et al.* (eds.), *Harvard Studies in Korean Linguistics* 3, 445-467.

Kim, Young-Joo. 1990. *The Syntax and Semantics of Korean Case: The Interaction between Lexical and Syntactic Levels of Representation.* Ph.D. thesis, Harvard University.

Kitahara, Hisatsugu. 1993. Inalienable possession constructions in Korean: Scrambling, the proper binding condition, and case-percolation. In Hajime Hoji and Patrica Clancy (eds.), *Japanese/Korean Linguistics* 2, 394-408.

Lee, Seong-yong. 2007. Two subject positions in multiple nominative constructions. *Journal of Language Sciences* 14, 239-262.

Lodge, Milton. 1981. *Magnitude Scaling: Quantitative Measurement of Opinions.* Beverley Hills, CA: Sage Publications.

Maling, Joan. and Soowon. Kim. 1992. Case assignment in the inalienable possession construction in Korean. *Journal of East Asian Linguistics* 1, 37-68.

Moon, Gui-Sun. 2000 The predication operation and multiple subject constructions in Korean: focusing on inalienable possessive constructions. *Studies in Generative Grammar* 10, 239-263.

Na, Younghee. and Geoffrey. Huck. 1993. On the status of certain island violations in Korean. *Linguistics and Philosophy* 16, 181-229.

O'Grady, William. 1991. *Categories and Case.* Amsterdam: John Benjamins Publishing.

Park, Byung-Soo. 2001. Constraints on multiple nominative constructions in Korean: A constraint-based lexicalist approach. *The Journal of Linguistic Science* 20, 147-190.

Park, H.-M. 2005. Constraint-based analyses on the Korean double nominative constructions. *The Journal of Studies in Language* 21, 87-111.

Park, Ki-Seong 1995. *The Semantics and Pragmatics of Case Marking in Korean: A Role and Reference Grammar Account.* Ph.D. thesis. The State University of New York at Buffalo.

Rhee, Seongha. 1999. On the Multiple Nominative Constructions in Korean. In Young-hwa Kim Il-kon Kim, and Jeong-won Park (eds.), *Linguistic Investigations*, 398-490. Seoul: Hankuk Publishing Co.

Ryu, Byong-Rae. 2010. Licensing nominals in the multiple nominative constructions in Korean: A mereological perspective. In Ryo Otoguro, Kiyoshi Ishikawa, Hiroshi Umemoto, Kei Yoshimoto, Yasunari Harada (eds.), *Proceedings of the 24th Pacific Asia Conference on Language, Information, and Computation* (PACLIC-24), 271-280

Ryu, Byong-Rae. 2013a. Multiple case marking as default case copying: A unified approach to multiple nominative and accusative constructions in Korean. The Korean Society for Language and Information Monthly Colloquium, May 18, 2013.

Ryu, Byong-Rae. 2013b. Multiple case marking as case copying: A unified approach to multiple nominative and accusative constructions in Korean. In

Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar* (HPSG-20), 182-202.

Ryu, Byong-Rae. 2013c. Multiple case marking constructions in Korean revisited. *Language and Information* 17.2, 1-27.

Schütze, Carson. 1996. Korean "case stacking" isn't: unifying non-case uses of case particles. In Kiyumi Kusumoto (ed.), *Proceedings of the 26th North East Linguistic Society* (NELS-26), 351-365.

Schütze, Carson. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology.* Chicago, IL: University of Chicago Press.

Schütze, Carson. 2001. On Korean "case stacking": The varied functions of the particles *-ka* and *-lul*. *The Linguistic Review* 18, 193-232.

Seo, Sang-Kyu. 2002. Compilation of the Korean National Corpus and its perspective. In Yun-pyo Hong (ed.), *Korean and Its Informationalization*, 255-292. Seoul: Taehaksa.

Stevenson, Stanley. 1975. *Psycholinguistics: Introduction to Its Perceptual, Neural, and Social Prospects.* New York: John Wiley.

Ura, Hiroyuki. 1996. *Multiple Feature-Checking: A Theory of Grammatical Function Splitting.* Ph. D. thesis. MIT.

Van Valin, Robert and Randy LaPolla. 1998. *Syntax: Structure, Meaning, and Function.* Cambridge: Cambridge University Press.

Van Valin, Robert. 2009. Case in Role and Reference Grammar. In Andrej Malchukov and Andrew Spencer (eds.), *The Oxford Handbook of Case*, 102-120. Oxford: Oxford University Press.

Van Valin, Robert and William. Foley. 1980. Role and reference grammar. In Edith Moravcsik and Jesica Wirth (eds.), *Current Approaches to Syntax*, 329-352. New York: Academic Press.

Yang, In-Seok. 1972. *Korean Syntax: Case Markers, Delimiters, Complementation, and Relativization.* Ph.D. thesis. University of Hawaii.

Yoon, Jeong-Me. 1997. The argument structure of relational nouns and inalienable possessor constructions in Korean. *Language Research* 33.2, 231-64.

Yoon, James. 1986. Some queries concerning the syntax of multiple subject constructions in Korean. *Studies in the Linguistic Sciences* 16, 215-236.

Submitted on: October 21, 2013
Revised on: December 8, 2013
Accepted on: December 16, 2013