Linguistic Research 35(3), 483-511
DOI: 10.17250/khisli.35.3.201812.003

Scrambling and island constraints in Korean: An experimental approach*

Yong-hun Lee
(Chungnam National University)

Lee, Yong-hun. 2018. Scrambling and island constraints in Korean: An experimental approach. *Linguistic Research* 35(3), 483-511. There have been several controversies on the existence of island constraints in Korean. This paper takes an experimental approach to island constraints in Korean and examines the properties of islandhood in Korean. This paper basically adopts the experimental designs in Sprouse et al. (2012, 2014), but one more factor *Scramble* is also considered in the experiments as in Kim and Goodall (2014). Although the latter study includes only two island constraints (*whether* island and adjunct island), this paper contains all of four island constraints in the former studies. The experiments are conducted with the magnitude estimation (ME) method, and the acceptability scores are measured with line drawings for one hundred students. After the experiment, the collected data are analyzed with three types of analyses: generalized linear model (GLM), DD scores, and random forests. Through the analysis, the followings are observed: (i) Korean also has island phenomena, (ii) scrambling increases the DD scores in all of four types of island constraints, and (iii) the increases of the DD scores are not due to the violation of island constraints but can be explained by a processing-based account. **(Chungnam National University)**

Keywords island constraints, scrambling, GLM, DD scores, random forests

1. Introduction

Since English island constraints were observed in Ross (1967), there have been a lot of debates and studies on the existence of island constraints in other languages. Some scholars claim that islands constraints are universal syntactic phenomena, but others mention that some languages demonstrate no island effects. There have also been much more debates on the status of island constraints in *wh*-in-situ languages (such as Japanese or Korean), since the island violations are not directly observed in these languages. For example, Nishigauchi

* This paper is an extension of Lee and Park (2015). I wish to thank two anonymous reviewers of this journal for their helpful comments and suggestions. All remaining errors, however, are mine.

(1990) and Watanabe (1992) claim that there exist island constraints in Japanese, but Ishihara (2002) and Sprouse et al. (2011) mention that this language has no island constraint. Likewise, there have been some controversies on the existence of island constraints in Korean. Some have argued for the presence of island effects (Lee 1982; Han 1992; Hong 2004), while others have argued against it (Sohn 1980; Kang 1986; Suh 1987; Hwang 2007).

The goal of this paper is to investigate how scrambling may influence island constraints in Korean. for this purpose, this paper takes an experimental approach and scrutinizes the island properties in this language. The research designs basically follow those in Sprouse et al. (2012, 2014), and the target sentences are constructed based on those in these studies. However, one more factor *Scrambling* is included in the experiments as in Kim and Goodall (2014). The difference between Kim and Goodall (2014) and the current study is that the former study includes only two types of island constraints (*whether* island and adjunct island) in the experiments but that the latter contains all of four island constraints (*whether* island, complex NP island, subject island, and adjunct island) as in Sprouse et al. (2012, 2014). The sentences in Sprouse et al. (2012, 2014) are converted into Korean counterparts. Then, the constructed target sentences are mixed with the fillers and the sentences are presented to the Korean native speakers. More than one hundred native speakers participated in the experiment, and their intuition is measured with a Magnitude Estimation (ME) method. After the experiment, all of the data for 100 participants are extracted and they are statistically analyzed with R. For these data, three types of analysis methods are adopted: GLM, DD scores, and random forests.

This paper is organized as follows. In Section 2, previous studies are reviewed especially focused on the experimental approaches. Section 3 includes the accounts for experimental designs (research materials and research methods) and procedures, and Section 4, Section 5, Section 6, and Section 7 enumerate the results of experiments. Section 8 contains discussions, and Section 8 summarizes this paper.

2. Previous studies

2.1. Theoretical approaches to island constraints in Korean

Since Ross (1967) identified the island constraints in English, there have been a lot of studies that investigate the presence or absence of island constraints in other languages. The focus of these previous studies primarily has been on examining (i) if the island constraints also exist in the languages of studies and (ii) why their languages escape the islands if the languages do not have the island constraints.

Korean is no exception, and there have been several studies on the existence of island constraints in Korean. Some earlier studies are primarily focused on the basic island properties in Korean. Choi (1989) explains the Korean island phenomena with LF-movements. That is, since Korean is a *wh*-in-situ language, the island constraints are hard to be observed in the S-structure. The covert *wh*-movements occur in the LF representation even in the *wh*-in-situ languages, and the island phenomena in the Korean island phenomena can be explained with LF-movements. Based on this idea, Song (1995) further investigates the relationship between the island constraints in Korean and *wh*-in-situ property. On the other hand, Lee (1999) studies whether the negative island constraint exists in Korean.

In syntactic literature, there are two opposite camps to island constraints in Korean. Some scholars claim that Korean has island constraints (Lee 1982; Han 1992; Hong 2004; Park 2001, 2009). Hong (2004) proposes island and intervention effects as two diagnostics for syntactic movements. The study says that Korean has island effects and that no intervention effects are observed in the *wh*-movements. Park (2001, 2009) examine various sluicing constructions in Korean. Through the investigation, it is found that matrix sluicing in Korean is island-sensitive. The study argues that the island sensitivity arises because *wh*-phrases do not move to [Spec, CP] in overt syntax. Park (2009) also provides accounts for the contrast between matrix sluicing and fragment answers in Korean with respect to island sensitivity.

On the other hand, other scholars claim that there is no island effect in Korean (Sohn 1980; Kang 1986; Suh 1987; Hwang 2007; Chung 2005; Yoon 2011,

2012; Kim 2013, 2017). For example, Chung (2005) mentions that Korean *ettehkey* (how) does not show any island effect. Given the revised nominal analysis, the scope of *ettehkey* (how) in Korean had to be licensed via binding, since there was no island effect. Yoon (2011, 2012) identifies two novel environments where *wh*-phrases show no island effects: the declarative intervention context and the embedded context. Then, the question is why the in-situ *wh*-phrases are not identical to the standard *wh*-phrases in English. The study also mentions that the standard *wh*-island effects correspond to the misinterpretation judgment and supports the claim by showing that there is a strong correlation between the *wh*-islands and the possibility that *wh*-in-situ questions would be misinterpreted as Yes/No-questions. Kim (2013) investigates *wh*-islands in the relative clauses. The study claims that the fact that the Korean language escape the island constraint can be explained by a complex semantico-pragmatic constraint, which is based on the notion of 'coherence' and the construction-specific factors which cause processing difficulty.

2.2. Experimental approaches to island constraints in English

Recently, as computer technology and statistics develop, many researchers try to study various syntactic phenomena through carefully-designed experiments and their statistical analysis. Some of them have had an interest in measuring native speakers' intuition on syntactic data objectively and scientifically (Bard et al. 1996; Schütze 1996; Cowart 1997; Keller 2000). This research method may also be applied to the study of islands, and lots of fruitful facts have been discovered through experimental approaches to syntax and their statistical analyses.

Sprouse et al. (2012) adopts an experimental approach to the island phenomena in English and examines native speakers' intuition on island constraints. They employ 2×2 factor combinations in (1) and investigate four types of island constraints using the following target sentences (Sprouse et al. 2012: 87-88).

(1) Factor Combinations
   a. Type 1: NON-ISLAND | MATRIX
   b. Type 2: NON-ISLAND | EMBEDDED

c. Type 3: ISLAND | MATRIX
d. Type 4: ISLAND | EMBEDDED

(2) *Whether* islands
a. Who __ thinks that John bought a car?
b. What do you think that John bought __ ?
c. Who __ wonders whether John bought a car?
d. What do you wonder whether John bought __ ?

(3) Complex NP islands
a. Who __ claimed that John bought a car?
b. What did you claim that John bought __?
c. Who __ made the claim that John bought a car?
d. What did you make the claim that John bought __?

(4) Subject islands
a. Who __ thinks the speech interrupted the TV show?
b. What do you think __ interrupted the TV show?
c. Who __ thinks the speech about global warming interrupted the TV show?
d. What do you think the speech about __ interrupted the TV show?

(5) Adjunct islands
a. Who __ thinks that John left his briefcase at the office?
b. What do you think that John left __ at the office?
c. Who __ laughs if John leaves his briefcase at the office?
d. What do you laugh if John leaves __ at the office?

Along with these sentences, they conduct an experiment and examine the intuition of 173 native speakers. Through the experiments, they obtain the following results (Sprouse et al. 2012: 100).

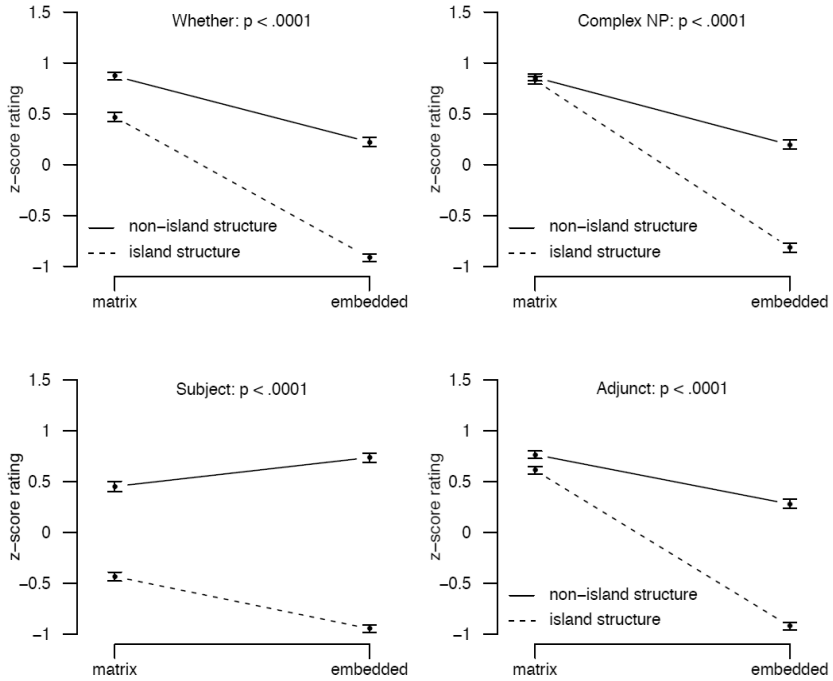

Figure 1. Analysis results in Sprouse et al. (2012)

These results illustrate (i) that native speakers show more acceptability for non-island structures than island structures both in matrix and embedded clauses and (ii) that the differences of acceptability scores become greater in embedded clauses than matrix clauses. These observations demonstrate that there are clearly island effects in English.

2.3. Experimental approaches to island constraints in Korean

Experimental approaches are also adopted in the studies of island constraints in Korean. Kim and Goodall (2014) employs a similar experimental method in Sprouse et al. (2012, 2014) to their experimental designs and examines the island constraints in Korean. They design four experiment sets in order to test the existence of two island constraints in Korean: *whether* island and adjunct island. Since Korean is a *wh*-in-situ language, they add another factor on word order

(canonical order vs. scrambled). Their experiments had a 2×2×2 design: Location of *wh*-word (matrix vs. embedded clause), Embedded clause type (non-island vs. island) and Answer type (appropriate for direct *wh*-question vs. yes/no question).

They utilize the question-answer pairs along with appropriate contexts in order to examine native speakers' intuition on island constraints. They make the questions in the stimuli ambiguous so that *wh*-words might be interpreted either as *wh*-words or as existential, as in Hong (2004).

A total of 48 native speakers participate in the experiments and the intuition is measured with a 7-point Likert scale. The following figures show us the analysis results.

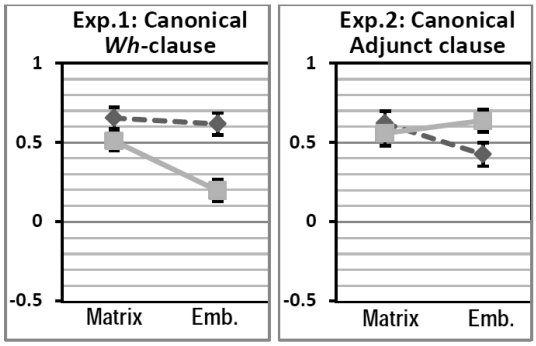

Figure 2. Analysis results in Kim and Goodall (2014): canonical order

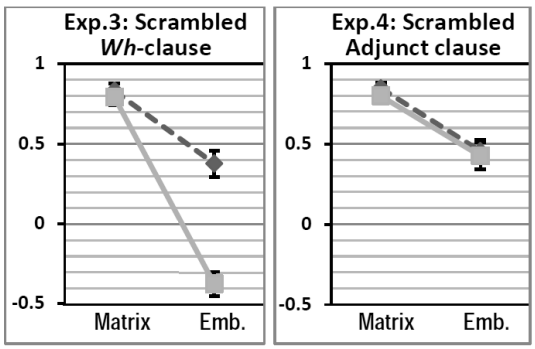

Figure 3. Analysis results in Kim and Goodall (2014): scrambled order

Through the experiment, they observe that there are significant interactions

between Location (matrix or embedded clause) and Embedded clause type (non-island or island) in Exp. 1 and Exp. 3, but that there is no such interaction in Exp. 2 and Exp. 4. This implies that there is an island effect with *wh*-clauses (Exp. 1 and Exp. 3) but that no island effect exists with adjunct clauses (Exp. 2 and Exp. 4).

Along with these results, they obtain another interesting observation. For the ambiguous questions contained in question-answer pairs, they find that one reading or the other reading is encouraged. Furthermore, they also observe that the presence or absence of an appropriate context crucially make the *wh*-reading pragmatically plausible or implausible, even in the cases where an island constraint is violated.

3. Research method

3.1. Research questions and hypothesis

This paper basically follows the experimental designs in Sprouse et al. (2012, 2014) and all of four types of island constraints in (2) - (5) are included in the experiments. However, since Korean is a language with free order, this paper also adopts the designs in Kim and Goodall (2014) and includes the experiments with two different options 'canonical' and 'scrambled'. The research questions are as follows.

(6) Research Questions
   a. Does Korean demonstrate the (four) island phenomena as in English?
   b. Does word order influence the acceptability scores of the island-related sentences in Korean?
   c. If word order does not affect the acceptability scores of the island-related sentences, which factor influences the acceptability scores?

For these questions, the following hypotheses are made.

(7) Hypotheses

a. If Korean really has island constraints, the differences between Type 2 and Type 4 in (1) will significantly be bigger than those between Type 1 and Type 3.

b. If word order affects the acceptability scores of the island-related sentences in Korean, there will be statistically significant differences between the DD scores of 'canonical' and those of 'scrambled'.

c. If word order does not affect the acceptability scores of the island-related sentences in Korean, other factor will increase or decrease the acceptability scores and the variable importance of the relevant factor will increase in the data set for 'scrambled' order.

These positive hypotheses are constructed for the research questions in (6a), (6b), and (6c) respectively, and the negative versions can be constructed by negating each sentence in (7).

3.2. Materials

In order to closely examine the island constraints in Korean, the first thing to be done is to make appropriate target sentences. This paper basically follows the factor combinations in (1) à la Sprouse et al. (2012, 2014), but another factor *Scramble* is also taken into consideration as in Kim and Goodall (2014). However, this factor is not directly reflected in the experimental design. Instead, the factor is reflected in the design by constructing two different sets of target sentences: 'canonical' order and 'scrambled' order. Accordingly, the following two factors are included in the actual experimental design: Island constraint (Absence vs. Presence) and Location of *wh*-word (Matrix clause vs. Embedded clause). Since two factors are employed and each factor has two values, the experiment has a 2×2 design.

First of all, basic target sentences were made with the sentences in Sprouse et al. (2012, 2014). For example, the following sentences are basic target sentences for Korean.

(8) *Whether* islands (canonical order)

a. ***Nwu-ka*** [*Chelsoo-ka* *ku* *cha-lul*
who.NOM [Chelsoo.NOM the car.ACC
sa-ss-ta-ko] *sayngkakha-ni*?
buy.PAST.DECL.COMP] think.Q
'Who thinks that Chelsoo bought the car?'

b. *Younghee-nun* [*Chelsoo-ka* ***mwues-ul***
Younghee.NOM [Chelsoo.NOM what.ACC
sa-ss-ta-ko] *sayngkakha-ni*?
buy.PAST.DECL.COMP] think.Q
'What does Younghee think that Chelsoo bought?'

c. ***Nwu-ka*** [*Chelsoo-ka* *ku* *cha-lul*
who.NOM [Chelsoo.NOM the car.ACC
sa-ss-nunci *an sa-ss-nunci*] *kungkumhayha-ni*?
buy.PAST.DECL.COMP not or] wonder.Q
'Who wonders that Chelsoo bought the car or not?'

d. *Younghee-nun* [*Chelsoo-ka* ***mwues-ul***
Younghee.NOM [Chelsoo.NOM what.ACC
sa-ss-nunci *an sa-ss-nunci*] *kungkumhayha-ni*?
buy.PAST.DECL.COMP not or] wonder.Q
'What does Younghee wonder whether Chelsoo bought or not?'

(9) Complex NP islands (canonical order)

a. ***Nwu-ka*** [*Chelsoo-ka* *ku* *cha-lul*
who.NOM [Chelsoo.NOM the car.ACC
sa-ss-ta-ko] *cucangha-ni*?
buy.PAST.DECL.COMP] claim.Q
'Who claims that Chelsoo bought the car?'

b. *Younghee-nun* [*Chelsoo-ka* ***mwues-ul***
Younghee.NOM [Chelsoo.NOM what.ACC
sa-ss-ta-ko] *cucangha-ni*?
buy.PAST.DECL.COMP] claim.Q
'What does Younghee claim that Chelsoo bought?'

c. ***Nwu-ka*** [*Chelsoo-ka* *ku* *cha-lul*
who.NOM [Chelsoo.NOM the car.ACC
sa-ss-ta-nun] *cucang-ul* *ha-ni*?
buy.PAST.DECL.COMP] claim.ACC do.Q
'Who makes the claim that Chelsoo bought the car?'

d. *Younghee-nun* [*Chelsoo-ka* ***mwues-ul***
Younghee.NOM [Chelsoo.NOM what.ACC
sa-ss-ta-nun] *cucang-ul* *ha-ni*?
buy.PAST.DECL.COMP] claim.ACC do.Q
'What does Younghee make the claim that Chelsoo bought?'

(10) Subject islands (canonical order)

a. ***Nwu-ka*** [*ku* *yensel-i* *ku* *hayngsa-lul*
who.NOM [the speech.NOM the event.ACC
panghayha-n-ta-ko] *sayngkakha-ni*?
interrupt.PRES.DECL.COMP] think.Q
'Who thinks that the speech interrupts the event?'

b. *Younghee-nun* [***mwues-i*** *ku* *hayngsa-lul*
Younghee.NOM [what.NOM the event.ACC
panghayha-n-ta-ko] *sayngkakha-ni*?
interrupt.PRES.DECL.COMP] think.Q
'What does Younghee think that interrupts the event?'

c. ***Nwu-ka*** [*ciku on.nanhwa-ey tayhan* *yensel-i*
who.NOM [global warming.about the speech.NOM
ku *hayngsa-lul* *panghayha-n-ta-ko*] *sayngkakha-ni*?
the event.ACC interrupt.PRES.DECL.COMP] think.Q
'Who thinks that the speech about global warming interrupts the event?'

d. *Younghee-nun* [*ciku* *on.nanhwa-ey tayhan* ***mwues-i***
Younghee.NOM [global warming.about what.NOM
ku *hayngsa-lul* *panghayha-n-ta-ko*] *sayngkakha-ni*?
the event.ACC interrupt.PRES.DECL.COMP] think.Q
'What does Younghee think that the speech about interrupts the event?'

(11) Adjunct islands (canonical order)

a. ***Nwu-ka*** [*Chelsoo-ka* *ku* *cha-lul* *samusil-ey*
who.NOM [Chelsoo.NOM the car.ACC office.LOC
tukow-ass-ta-ko] *sayngkakha-ni*?
leave.PAST.DECL.COMP] think.Q
'Who thinks that Chelsoo left his book at the office?'

b. *Younghee-nun* [*Chelsoo-ka* ***mwues-ul*** *samusil-ey*
Younghee.NOM [Chelsoo.NOM what.ACC office.LOC
tukow-ass-ta-ko] *sayngkakha-ni*?
leave.PAST.DECL.COMP] think.Q
'What does Younghee think that Chelsoo left at the office?'

c. ***Nwu-ka*** [*Chelsoo-ka* *ku* *cha-lul* *samusil-ey*
who.NOM [Chelsoo.NOM the car.ACC office.LOC
tukow-ass-ta-myen] *piwuskeyss-ni*?
leave.PAST.DECL.COMP] laugh.Q
'Who laughs if Chelsoo left his book at the office?'

d. *Younghee-nun* [*Chelsoo-ka* ***mwues-ul*** *samusil-ey*
Younghee.NOM [Chelsoo.NOM what.ACC office.LOC
tukow-ass-ta-myen] *piwuskeyss-ni*?
leave.PAST.DECL.COMP] laugh.Q
'What does Younghee laugh if Chelsoo left at the office?'

These sentences are for the 'canonical' word order. In the sentence for 'scrambled' word order, the bracketed part (the [] part) in each sentence moves to the sentence-initial position.

Four target sentences are constructed more by changing the lexical items in each island constraint. A total of twenty target sentences are constructed for each island constraint. Along with these target sentences, the double number of filler sentences are made for each island constraint. The half of the filler sentences (20 sentences) are constructed based on the structure of the target items. However, they are not related with the corresponding island constraint. The others of the filler sentences (20 sentences) are composed of the sentences that have no relation with the purpose of the experiments. Among them, 10 sentences are grammatical ones and the others are ungrammatical ones.

After all the target and filler sentences are constructed, a set of random numbers is generated with the R function, and each sentence is given the generated random numbers. Then, the sentences are given to the participants after the sentences are sorted based on the random number.

3.3. Procedures

The experiments are performed in the on-line survey forms, which is constructed by SurevyGizmo.[1] After the experiment, the data for a total of 100 native speakers are collected through the filtering process (m=21.79, sd=2.342).[2] All the participants reside in and around Daejeon area, South Korea. They are either current university students or graduates of universities in Korea. The experiments are approved by the Institutional Review Board (IRB) of the Hannam University (#17-04-01-0201). All subjects involved gave their informed written consent.

All the participants are first asked to fill out a simple one-page survey that contains biographical information such as age, gender, and dialect(s). Then, after a short exercise using some example sentences, they are asked to proceed to take the main task. The main task is composed of acceptability judgment tasks using the Magnitude Estimation (ME) (Lodge 1981; Johnson 2008).[3] Since the *wh*-words in Korean can be interpreted either as *wh*-phrases or as indefinite pronouns, the relevant context is given to every sentence so that all the *wh*-words can be interpreted as *wh*-phrases (not as indefinite pronouns). In addition, following Sprouse (2008), the modulus sentences are provided in every page against which the participants measure the acceptability scores.

Actually, there are two types of ME methods: numerical estimates and line drawing. Bard et al. (1996) pointed out that the participants sometimes think of

---

1 http://www.surveygizmo.com

2 In the filtering process, the data for some participants are excluded in the cases where not all the sentences are answered or where the scores belong to outliers.

3 There are several reasons why this paper took an ME in the acceptability judgment task, rather than the Likert scale. Lee (2013) contained a detailed discussion on the differences between ME and Likert scales in the acceptability judgment task (intuition tests). Lodge (1981) mentioned that this ME had several advantages over the category scaling (the Likert scale). Although there are some claims that the Likert scales are available in the acceptability judgment task, this paper follows previous studies (Lodge 1981; Johnson 2008) and adopted ME in the experiment.

numeric estimates as something like academic test scores and that 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 ME. Accordingly, the current study adopts a line drawing method in which the participants are asked to draw different lengths of lines to indicate the naturalness (acceptability) of a given sentence (after reading the sentence). An acceptability judgment task (also known as native speakers' intuition test) is used in the study since this method is known to be a psychological experiment which can be used to get the subconscious knowledge of native speakers in a given language (Carnie 2012). In the main task, participants are required to draw a line for each sentence, according to the degree of acceptability/naturalness of the given sentence.

3.4. Statistical analysis

All of the statistical analyses in this paper are performed by R (R Core Team, 2017). After all the data are collected from acceptability judgment tasks, the scores are extracted for target sentences by measuring the length of lines, and the scores are normalized with *z*-scores. Then, the normality tests (Baayen 2008; Gries 2013) are performed to check whether parametric tests are available or not. If the distributions of the data follow the normal distribution, the parametric tests are available, such as *t*-tests, ANOVAs, or (ordinary) linear regression tests. However, if the distributions do not follow the normal distribution, the non-parametric tests must be applied such as Wilcoxon tests, Friedman tests, or generalized linear regression tests.

When the normality tests are performed, it is found that all the data sets do not follow the normal distribution. Some data sets are positively skewed, and other sets showed a slightly bimodal distribution. Accordingly, non-parametric tests have to be applied in the analysis of the collected data.

After the normality tests, the collected data are descriptively analyzed. Then, in order to closely examine how each factor affects the acceptability of the target sentences, a Generalized Linear (Regression) Model (GLM) is performed. According to Agresti (2007), a GLM is available when the distribution does not follow the normal distribution. Thus, the test is adopted to examine how each factor affects the acceptability of the sentences.

4. Analysis results: Canonical orders

The first data set is for the 'canonical' order, where the bracketed parts are located in the base position (i.e., non-scrambled position). The line plots in Figure 4 demonstrate the analysis results.

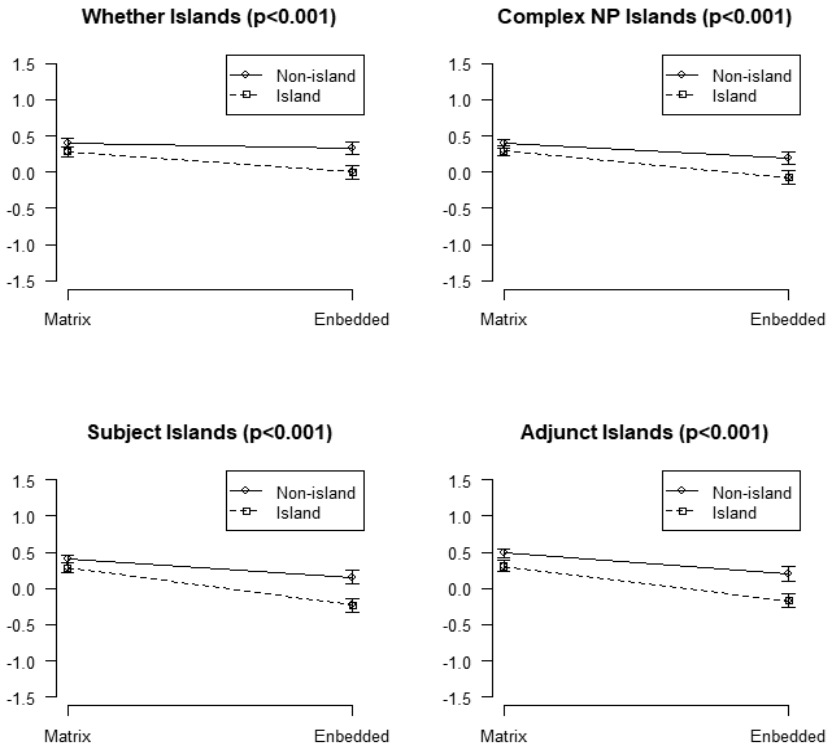

Figure 4. Analysis results (canonical order)

Although the acceptability scores in the 'Island' sentences seem bigger than those in 'Non-island' sentences, we cannot be sure if the changes are statistically significant. Accordingly, the GLM analyses are conducted for each island constraint, and the following four tables illustrate the analysis results.

Table 1. GLM analysis for *Whether* islands (canonical order)

	Estimate	sd	t	p	
(intercept)	0.254	0.020	12.789	<.001	***
Clause	-0.091	0.020	-4.573	<.001	***
Island	-0.114	0.020	-5.717	<.001	***
Clause:Island	-0.053	0.020	-2.650	.008	**

Table 2. GLM analysis for complex NP islands (canonical order)

	Estimate	*sd*	*t*	*p*	
(intercept)	0.202	0.020	10.008	<.001	***
Clause	-0.141	0.020	-6.962	<.001	***
Island	-0.094	0.020	-4.628	<.001	***
Clause:Island	-0.043	0.020	-2.126	.034	*

Table 3. GLM analysis for subject islands (canonical order)

	Estimate	*sd*	*t*	*p*	
(intercept)	0.153	0.020	7.448	<.001	***
Clause	-0.192	0.020	-9.379	<.001	***
Island	-0.128	0.020	-6.224	<.001	***
Clause:Island	-0.066	0.020	-3.212	.001	**

Table 4. GLM analysis for adjunct islands (canonical order)

	Estimate	*sd*	*t*	*p*	
(intercept)	0.206	0.021	9.680	<.001	***
Clause	-0.190	0.021	-8.956	<.001	***
Island	-0.138	0.021	-6.492	<.001	***
Clause:Island	-0.048	0.021	-2.271	.023	*

As you can see in these tables, two factors *Clause* and *Island* significantly affect the acceptability scores. It implies that not only the distinction between 'Matrix' and 'Embedded' but also the distinctions between 'Island' and 'Non-island' plays a crucial role in the determination of acceptability score in the target sentences. Also note that there are some interactions between these two factors (i.e., *Clause*:*Island*) and that the interactions are statistically significant ($p<0.05$) for all of four types of island constraints.

5. Analysis results: Scrambled orders

The second data set is for the 'scrambled' order, where the bracketed parts are located in the sentence-initial position. The line plots in Figure 5 show us the analysis results.

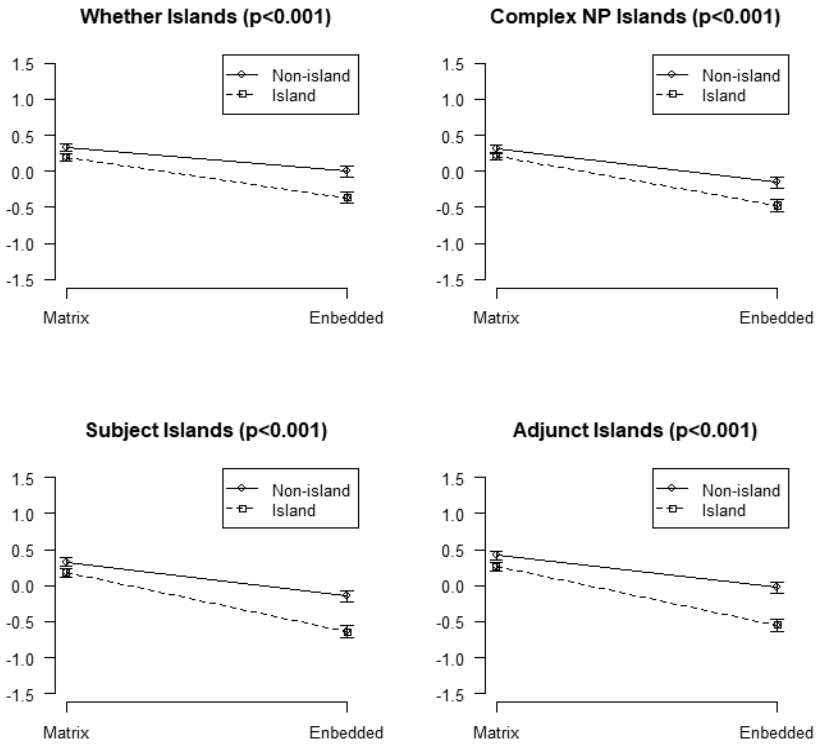

Figure 5. Analysis results (scrambled order)

Unlike the line plots in Figure 4, the differences of acceptability scores in the 'Island' sentences seem much bigger than those in 'Non-island' sentences. To examine if the changes are statistically significant, the GLM analyses are conducted once again for each island constraint, and the following four tables present the analysis results.

Table 5. GLM analysis for *Whether* islands (scrambled order)

	Estimate	*sd*	*t*	*p*	
(intercept)	0.038	0.017	2.251	.024	*
Clause	-0.222	0.017	-13.048	<.001	***
Island	-0.125	0.017	-7.328	<.001	***
Clause:Island	-0.058	0.017	-3.402	<.001	***

Table 6. GLM analysis for complex NP islands (scrambled order)

	Estimate	*sd*	*t*	*p*	
(intercept)	-0.027	0.018	-1.515	.130	
Clause	-0.288	0.018	-16.129	<.001	***
Island	-0.106	0.018	-5.926	<.001	***
Clause:Island	-0.057	0.018	-3.184	.001	**

Table 7. GLM analysis for subject islands (scrambled order)

	Estimate	*sd*	*t*	*p*	
(intercept)	-0.070	0.018	-3.861	<.001	***
Clause	-0.324	0.018	-17.967	<.001	***
Island	-0.160	0.018	-8.867	<.001	***
Clause:Island	-0.089	0.018	-4.912	<.001	***

Table 8. GLM analysis for adjunct islands (scrambled order)

	Estimate	*sd*	*t*	*p*	
(intercept)	0.027	0.018	-1.468	.142	
Clause	-0.314	0.018	-16.978	<.001	***
Island	-0.168	0.018	-9.096	<.001	***
Clause:Island	-0.091	0.018	-4.905	<.001	***

As in these tables of Section 4, two factors *Clause* and *Island* significantly affect the acceptability scores. It implies that not only *Clause* but also *Island* plays a significant role in the determination of acceptability score in the target sentences. Also note that, as in the sentences for 'canonical' orders, there are some interactions between these two factors (i.e., *Clause*:*Island*) and that the interactions are statistically significant (p<0.05) in all types of island constraints.

6. Analysis with DD-scores

In addition to statistical analyses, the differences-in-differences (DD) scores can be used to measure the strength of island constraints (Maxwell and Delaney 2003). The DD scores are calculated as follows.

(12) Calculation of DD scores
a. D1 = NON-ISLAND|EMBEDDED − ISLAND|EMBEDDED
b. D2 = NON-ISLAND|MATRIX − ISLAND|MATRIX
c. DD = D1 − D2

Sprouse et al. (2012, 2014) also utilize1 the DD scores to investigate the island phenomena in English. Following the traditions, this paper also includes the analysis of DD scores for Korean data. Table 9 and Table 10 illustrate the analysis results.

Table 9. DD scores (canonical order)

	Type 1	Type 2	Type 3	Type 4	D1	D2	DD
Whether	0.406	0.328	0.284	-0.003	0.331	0.122	0.209
Complex NP	0.393	0.197	0.292	-0.075	0.272	0.101	0.171
Subject	0.407	0.154	0.283	-0.233	0.387	0.124	0.263
Adjunct	0.485	0.202	0.306	-0.171	0.373	0.179	0.194

Table 10. DD scores (scrambled order)

	Type 1	Type 2	Type 3	Type 4	D1	D2	DD
Whether	0.327	-0.001	0.193	-0.366	0.365	0.134	0.231
Complex NP	0.309	-0.152	0.212	-0.478	0.326	0.097	0.229
Subject	0.325	-0.146	0.184	-0.643	0.497	0.141	0.356
Adjunct	0.417	-0.027	0.263	-0.545	0.518	0.154	0.364

According to previous studies including Sprouse et al. (2012, 2014), we can say that a given language has island effects if the DD scores have plus values. As Table 9 and Table 10 clearly demonstrate, the DD scores have plus values for all of the island constraints. If we follow the traditions of previous studies, we cannot help saying that Korean also has island constraints.

In order to examine if the DD scores in 'scrambled' order significantly higher than those of 'canonical' order, a Wilcoxon signed rank test (the non-parametric version of paired *t*-test) is conducted for each island constraint. The result is that there are statistically significant differences in all of the island constraints (*Whether*: V=58159, p=0.039; Complex NP: V=71659, p<0.001; Subject: V=65414, p=0.023; Adjunct: V=78559, p<0.001).

7. Random forests analysis

Though it is found that there are statistically significant increases of the DD scores in 'scrambled' order, compared with those of 'canonical' order (in all of the island constraints), it is necessary to examine where the differences are originated from. Do the differences of DD scores come from the factor *Island*? In order to answer this question, random forests analyses are conducted for 'canonical' order and 'scrambled' order respectively.

A *random forest analysis* or a random decision forest analysis is both a statistical method and a machine learning method, which can be used for various types of statistical tasks including classification and regression. This analysis usually operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Breiman 2001; Hastie et al. 2008).

One of the advantages of adopting a random forest analysis in linguistic analyses is that the variable importance can be calculated for each (linguistic) factor/variable. As its name indicates, the analysis *randomly* extracts the subset from the original data set and performs the analyses. Since the extractions are conducted several times, the subsets of each extraction form a *forest*. In this paper, random forest analyses are performed a thousand times, and variable importance is calculated at each time.

Since the factor *Island* is the main focus in this paper, the distributions of (absolute) variable importance (1,000 values) are calculated for 'canonical' order and 'scrambled' order respectively, and they are represented in Figure 6.

As you can find, the distribution of variable importance (of the factor *Island*)

for 'canonical' order clearly overlaps with that of 'scrambled' order. For the purpose of clarifying the fact that two distributions are not significantly different, a paired *t*-test is conducted. The result is that two distributions are not significantly different (t=-0.969, p=0.332).[4] It implies that the variable importance of *Island* does not increase significantly in the data set for 'scrambled' order.

Then, why do the DD scores increase significantly in the data set for 'scrambled' order? To provide a possible answer, the variable importance of all the factors are measured and shown in Figure 7.

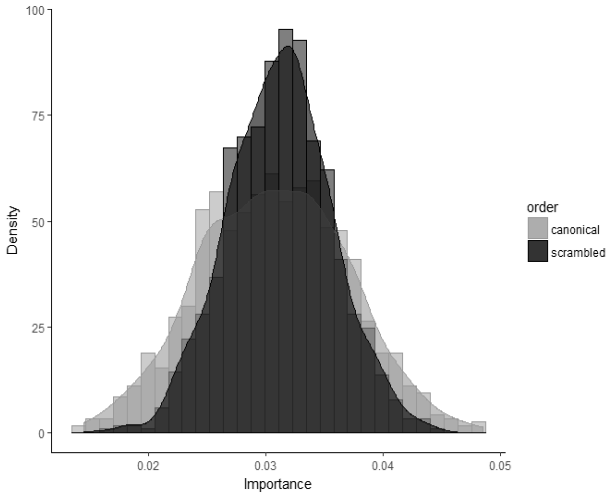

Figure 6. Analysis results (random forest, *Island*)

[4] The reason why a paired *t*-test is used here is that each distribution follows the normal distribution.

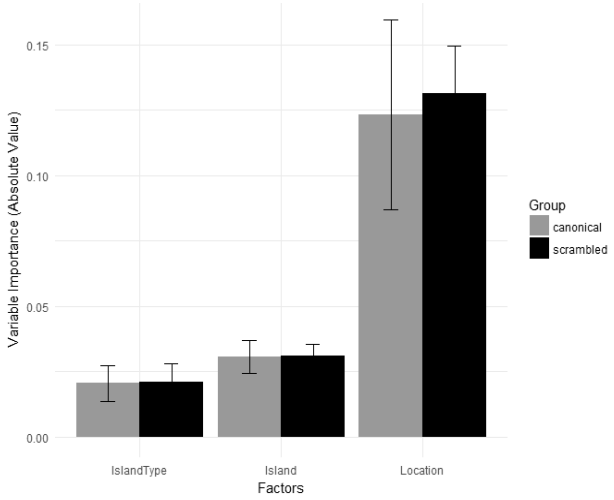

Figure 7. Analysis results (random forest, all factors)

Here, *IslandType* refers to four island types in the experiments (*Whether*, Complex NP, Subject, and Adjunct). As you can identify, the variable importance of *Location* (Matrix vs. Embedded) increases in the data set for 'scrambled' order. Then, there is a possibility that this factor significantly increases the DD scores in the data set for 'scrambled' order.

8. Discussions

There are several and continuous controversies on the existence of island constraints in Korean. Some scholars claim that Korean also has island constraints (Lee 1982; Han 1992; Hong 2004; Park 2001, 2009) and others say that Korean does not show island phenomena (Sohn 1980; Kang 1986; Suh 1987; Chung 2005; Hwang 2007; Yoon 2011, 2012; Kim 2013, 2017).

In order to examine the island status of Korean, this paper adopts the experimental designs in Sprouse et al. (2012, 2014) and investigates the island phenomena in Korean. This paper basically uses the Korean translations of English sentences in Sprouse et al. (2012, 2014). However, the experimental design of this paper is different from previous studies in following two facets.

First, as in Kim and Goodall (2015), this paper includes one more factor *Scramble*. That is, the target sentences are constructed for two data sets: 'canonical' order and 'scrambled' order. Second, unlike most of previous studies, this paper adopted the ME method (Lodge 1981; Johnson 2008). The collected data are statistically analyzed with two different methods: GLM and DD-scores.

The analysis results in Figure 4 seems to support the claim that Korean does not show island phenomena (Sohn 1980; Kang 1986; Suh 1987; Chung 2005; Hwang 2007; Yoon 2011, 2012; Kim 2013), since there is little difference between the acceptability scores of 'Non-island' sentences and those of 'Island' sentences. However, the GLM analyses in Section 4 show that the differences are statistically significant. This implies that Korean may have island phenomena, supporting the previous claims that Korean has island constraints (Lee, 1982; Han, 1992; Hong, 2004; Park, 2001, 2009). Through the analysis, it is found that the factor *Island* plays an important role in the determination of acceptability scores of island-violating sentences in Korean and that this factor also interacts with another factor *Clause*. These tendencies occur in all of four types of island constraints. If we accept that previous studies of utilizing the DD scores for examining islandhood of a language are on the right track, we cannot help admitting that Korean has island phenomena, because plus values are observed in all the types of island constraints.

The same tendencies are observed in the data sets for 'scrambled' word order. That is, the GLM analyses show that the difference between the acceptability scores of 'Non-island' sentences and those of 'Island' sentences are statistically significant and that the factor *Island* also interacts with another factor *Clause*. The analyses of DD scores also demonstrated Korean has island phenomena. These tendencies occur in all of four types of island constraints.

An interesting fact is that the DD scores increase in the 'scrambled' order and the increases in DD scores are statistically significant. This observation implies that the factor *Island* may drop the acceptability scores further in the sentences with 'scrambled' order. However, random forest analyses in Section 7 indicate that the increases in DD scores are due to the factor *Island* (Figure 6), since the variable importance of this factor does not change at all in the data set for 'scrambled' order. In fact, it is a natural result. Even though the bracketed parts are scrambled into the sentence-initial position in the sentences of (8) - (11)

in the data set for 'scrambled' order, the *wh*-phrases do not escape the island domain. They will move to the [Spec, CP] position by covert LF movements. Thus, there is no reason to increase the variable importance of the factor *Island* even in the data set for 'scrambled' order.

Instead, the analysis results in Figure 7 opens the possibility that the factor *Location* increases the DD scores in the data set for 'scrambled' order. You may wonder whether this possibility is available or not, but a processing-based account can provide an answer. The basic ideas of processing-based accounts (Frazier and Clifton 1989; Kluender and Kutas 1993; Kluender 1998; Hofmeister and Sag 2010) are that human beings have limited working memory capacity and that the acceptability scores of sentences becomes lower when the processing cost exceeds the working memory capacity.

It is possible to apply these basic ideas to our data sets. The sentences with 'scrambled' order may require some portions of working memory capacity for processing scrambling, since the 'scrambled' part must be re-located in the original location in the semantic interpretation. Since human beings have limited working memory capacity and some portions of working memory capacity are necessary for handling scrambling, the available memory capacity may be reduced when Korean native speakers process the 'scrambled' sentences. Along with this reduced memory capacity, Korean native speakers have to process the other syntactic parts of island- related sentences. In the 'scrambled' sentences, they have to identify where the *wh*-phrases are located: whether they are originated from the matrix clause or from the embedded clause. Consequently, the sentence processing in the 'scrambled' sentences requires more burden on working memory capacity, and the burden may cause the drop in the acceptability scores (Type 2 and Type 4) in the 'scrambled' sentences. Therefore, that the factor *Location* increases the DD scores in the data set for 'scrambled' order. However, the significant change of variable importance does not necessarily imply the significant differences of DD scores. A small increase in variable importance may result in significant change in the DD scores.

The analysis results in this paper are also different from those in Kim and Goodall (2014). As Figure 2/3 and Figure 4/5 demonstrate, similar tendencies are observed in the *Whether* islands. However, the Adjunct islands are different. The island constraints are not observed in Kim and Goodall's study (2014), but the

island phenomena are observed in this paper. There may be some possible reasons for this discrepancies, including the following two. First, the experimental design is different. Kim and Goodall (2014) employs question-answer pairs while this paper adopts contexts and modulus sentences. This difference of experimental design may result in the differences in analysis results. Second, while Kim and Goodall (2014) uses Likert scales whereas this paper utilizes the ME method. Since the ME methods can capture more fine-grained differences (Bard et al. 1996; Schütze 1996; Cowart 1997; Keller 2000), the ME methods in this paper might be more sensitive to more fine-grained distinctions. Of course, more studies are necessary to identify why the analysis results are different between the experiments in Kim and Goodall (2014) and those in this paper.

Returning to our research questions and hypotheses in Section 3.1, what answers do the analysis results provide to them? The first question is if Korean demonstrates the (four) island phenomena as in English. The hypothesis in (7a) says that if Korean really has island constraints, the differences between Type 2 and Type 4 in (1) will significantly be bigger than those between Type 1 and Type 3. Figure 4/5 and the DD scores in Table 9/10 illustrate that the hypothesis in (7a) can be supported. Thus, we can provide an answer 'yes' to the first question. The second question is if *Scrambling* affects the acceptability scores of the island-related sentences in Korean. The hypothesis in (7b) says that there will be statistically significant differences between the DD scores of 'canonical' and those of 'scrambled' if word order affects the acceptability scores of the island-violating sentences in Korean. The comparisons of Figure 4/5 and the statistical test of DD scores in Table 9/10 illustrate that the hypothesis in (7b) can be supported, and we can provide an answer 'yes' to the second question. The third question is if word order does not affect the acceptability scores of the island sentences, which factor influences the acceptability scores. The hypothesis in (7c) mentions that other factor will increase or decrease the acceptability scores and the variable importance of the relevant factor will increase in the data set for 'scrambled' order, if word order does not affect the acceptability scores of the island-violating sentences in Korean. The analysis results of random forests demonstrate that the variable importance of *Island* does not increase significantly in the data set for 'scrambled' order but that there is a

possibility that this factor *Location* significantly increases the DD scores in the data set for 'scrambled' order. This implies that the hypothesis in (7c) is partially supported.

9. Conclusion

In this paper, the status of island constraints is closely examined by experimental method and its statistical analysis. Following the experimental design in Sprouse et al. (2012, 2014), the experiments are designed. However, the experimental design of this paper is different from previous studies, since this paper includes one more factor *Scramble* as in Kim and Goodall (2014) and this paper adopted the ME method (Lodge 1981; Johnson 2008). The target sentences are constructed based on the English sentences in Sprouse et al. (2012, 2014), and data sets are constructed both for 'canonical' order and 'scrambled' order. Then, the collected data are statistically analyzed using R with three different methods: GLM, DD scores, and random forests.

The analysis results demonstrate that Korean has island phenomena and that the island constraints occur in all of four types of island constraints. In addition, the comparisons of the analysis results in Section 4/5 and the comparison of DD scores in Table 9/10 illustrate that the DD scores significantly increase in the data set for 'scrambled' order. However, the analysis results of random forests indicate that the variable importance of *Island* does not change in the data set for 'scrambled' order. Instead, the result opens the possibility that the factor *Location* significantly increases the DD scores in the data set for 'scrambled' order. I hope that the analysis results in this paper can give us more opportunities to understand the status of island constraints in Korean.

References

Agresti, Alan. 2007. *An introduction to categorical data analysis*. 2nd edition. Hoboken, NJ: John Wiley & Sons.

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.
Breiman, Leo. 2001. Random forests. *Machine Learning* 45(1): 5-32.
Carnie, Andrew. 2012. *Syntax*: *A generative introduction*. 3rd Edition. Oxford: Blackwell.
Choi, Kiyong. 1989. LF movement: The *wh*-island Constraint. *Harvard Studies in Korean Linguistics* 3: 213-234.
Chung, Daeho. 2005. Why is HOW in Korean insensitive to islands?: A revised nominal analysis. *Studies in Modern Grammar* 39: 115-131.
Cowart, Wayne. 1997. *Experimental syntax*: *Applying objective methods to sentence judgments*. Thousands Oaks, CA: Sage Publications.
Frazier, Lyn and Charles Clifton Jr. 2002. Processing 'd-linked' phrases. *Journal of Psycholinguistic Research* 31: 633-660.
Han, Jong-Im. 1992. Syntactic movement analysis of Korean relativization. *Language Research* 28: 335-357.
Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2008. *The elements of statistical learning* (2nd edition). Berlin: Springer.
Hofmeister, Philip and Ivan Sag. 2010. Cognitive constraints on syntactic islands. *Language* 86: 366-415.
Hong, Sun-Ho. 2004. On the lack of syntactic effects in Korean *wh*-questions. *The Linguistic Association of Korea Journal* 12(3): 43-57.
Hwang, Heeju. 2007. *Wh*-phrase questions and prosody in Korean. In *Proceedings of 17th Japanese/Korean Linguistics*. Stanford, CA: CSLI.
Ishihara, Shinichiro. 2002. Invisible but audible *wh*-scope marking: *Wh*-constructions and deaccenting in Japanese. In Line Mikkelsen and Christopher Potts (eds.), *Proceedings of the Twenty-first West Coast Conference on Formal Linguistics*, 180-193. Somerville, MA: Cascadilla Press.
Johnson, Keith. 2008. *Quantitative methods in linguistics*. Oxford: Blackwell.
Kang, Young-Se. 1986. *Korean syntax and universal grammar*. PhD Dissertation. Harvard University.
Keller, Frank. 2000. *Gradient in grammar*: *Experimental and computational aspects of degrees of grammaticality*. PhD Dissertation. University of Edinburgh.
Kim, Boyoung and Grant Goodall. 2014. An experimental investigation of island effects in Korean. *Japanese/Korean Linguistics* 23. Stanford, CA: CSLI.
Kim, Ilkyu. 2013. Rethinking 'island effects' in Korean relativization. *Language Sciences* 38: 59-82.
Kim, Ilkyu. 2017. An experimental study on island effects related to 'double relative clauses' in Korean. *Linguistic Research* 34(3): 191-214.
Kluender, Robert and Marta Kutas. 1993. Subjacency as a processing phenomenon.

Language and Cognitive Processes 8: 573-633.

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. *Syntax and Semantics* 29: 241-279.

Lee, Doo-Won. 1999. Remarks on negative islands in Korean. *Harvard Studies in Korean Linguistics* 8: 457-471.

Lee, Hyo-Sang. 1982. Asymmetry in island constrains in Korean, ms., University of California at Los Angeles.

Lee, Yong-Hun and Yeonkyung Park. 2015. Complex-NP islands in Korean: An experimental approach. In *Proceedings of 29th Pacific Asia conference on language, information and computation* (PACLIC-29), 244-253. Shanghai, China: Jiao Tong University.

Lee, Yong-hun. 2013. Experimental approach to multiple Case constructions in Korean. *Language and Information* 17(2): 29-50.

Lodge, Milton. 1981. *Magnitude scaling*: *Quantitative measurement of opinions*. Beverley Hills, CA: Sage Publications.

Maxwell, Scott and Harold Delaney. 2003. *Designing experiments and analyzing data*: *A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer Academic Publishers.

Park, Bum-Sik. 2001. Island-insensitive sluicing. *Harvard Studies in Korean Linguistics* 9: 669-682.

Park, Bum-Sik. 2009. Island sensitivity in ellipsis and its implications for movement. *Studies in Generative Grammar* 19(4): 599-620.

Pearl, Lisa and Jon Sprouse. 2014. Computational models of acquisition for island. In Jon Sprouse and Norbert Hornstein (eds.), *Experimental syntax and island effects*, 109-131. Cambridge: Cambridge University Press.

R Core Team. 2017. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Ross, John. 1967. *Constraints on variables in syntax*. PhD Dissertation. Massachusetts Institute of Technology.

Schütze, Carson. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago, IL: University of Chicago Press.

Sohn, Ho-Min. 1980. Theme-prominence in Korean. *Korean Linguistics* 2: 2-19.

Song, Jae-Gyun. 1995. *Wh*-in-situ and island phenomena in Korean. *Harvard Studies in Korean Linguistics* 6: 402-412.

Sprouse, Jon, Mathews Wagers, and Colin Phillips. 2012. A test of the relation between working memory capacity and syntactic island effects. *Language* 88: 82-123

Sprouse, Jon, Matthew Wagers, and Colin Phillips. 2014. Deriving competing predictions from grammatical approaches and reductionist approaches to island effects. In Jon Sprouse and Norbert Hornstein (eds.), *Experimental syntax and island effects*, 21-41.

Cambridge: Cambridge University Press.
Sprouse, Jon, Shin Fukuda, Hajime Ono, and Robert Kluender. 2011. Reverse island effects and the backward search for a licensor in multiple *wh*-questions. *Syntax* 14: 179-203.
Sprouse, Jon. 2008. Magnitude estimation and the non-linearity of acceptability judgments. In Natasha Abner and Jason Bishop (eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 397-403. Somerville, MA: Cascadilla Proceedings Project.
Suh, Chungmok. 1987. Wh-*constructions in Korean*. Seoul: Top Press.
Watanabe, Akira. 1992. *Wh*-in-situ, subjacency, and chain formation. In *MIT Occasional Papers in Linguistics 2*. Cambridge, MA: MIT Press.
Yoon, Jeong-Me. 2011. *Wh*-island effects of *wh*-in-situ questions in Korean. *Studies in Generative Grammar* 21(4): 763-778.
Yoon, Jeong-Me. 2012. *Wh*-island effects in Korean *wh*-in-situ questions. *Korean Journal of Linguistics* 37(2): 357-382.

Yong-hun Lee
Lecturer
Department of English Language & Literature
Chungnam National University
99 Daehak-ro, Yuseng-gu,
Daejeon 34134, Korea
E-mail: yleeuiuc@hanmail.net

Received: 2018. 05. 01.
Revised: 2018. 07. 04.
Accepted: 2018. 10. 09.