English Island Constraints by Natives and Korean Non-natives

Yong-hun Lee* and Yeonkyung Park**
Chungnam National University
Hannam University

*First Author / **Corresponding Author

ABSTRACT

This paper examined how native English speakers (NS) and non-native Korean speakers (NNS) processed English island sentences differently, by taking an experimental design and its statistical analysis. Sprouse et al. (2012) demonstrated that NS identified four types of island sentences (Whether, Complex NP, Subject, and Adjunct) in their language. Kim (2015), on the other hand, adopted the similar method and observed that Korean NNS did not identify two island constraints (Complex NP and Adjunct islands). This study performed similar experiments in previous studies but it adopted the magnitude estimation (ME) method, rather than the Likert scale. The target sentences came from Sprouse et al. (2012), and the same set of sentences were used for both American NS and Korean NNS. A total of 100 students participated in the experiment respectively, and the acceptability scores were gauged with the ME method. Then, the collected data were analyzed with a generalized linear model (GLM). Through the experiments, the followings were found: (i) Korean NNS identified all of the four types of island constraints in English and (ii) DD-scores of Korean NNS were smaller than those of American NS. (Chungnam National University · Hannam University)

Keywords: island constraints, American native speakers, Korean non-native speakers, magnitude estimation, generalized linear model

1. Introduction

An island means a syntactic domain from which a movement operation is not permissible. Since Ross’s identification of island constraints in English (1967), there have been lots of debates on the status of the island constraints in many natural languages, including English. If all the natural languages contain island phenomena, they must be included in the Universal Grammar (UG). If not, the island constraints are language-specific phenomena. Some languages have been said that they have the island constraints, whereas other languages (e.g. Chinese, Korean, or Japanese) do not demonstrate island phenomena.

In addition to the continuous interest in theoretical syntax, the island constraints have also been discussed in the literature of Second Language Acquisition (SLA).
The studies in the SLA literature were motivated to investigate the possibility that (adult) L2 learners illustrated sensitivity to island constrains in the target language L2 (Hawkins and Hattori 2006, Johnson & Newport 1991, Schachter 1990, White and Genesee 1996, White and Juffs 1998, and many others), since the island effects of the L1 (the mother tongue) surely might influence the acquisition of L2 (L1 transfer effects; Selinker 1969, Odlin 1989, Odlin 2003).

Although there were a lot of theoretical studies on island constraints in English, Sprouse et al. (2012) took an experimental approach and examined the native speakers’ intuition to the island constraints. The study tried to check if NS clearly identified four types of islands in English: Whether, Complex NP, Subject, and Adjunct islands. On the other hand, Kim (2015) applied the similar method in Sprouse et al. (2012) and tested the islandhood of the English island sentences by the Korean NNS. Through the method, she found that Korean NNS clearly identified two island constraints (Whether islands and the Subject islands) but that they did not identify the other twos (Complex NP island and the Adjunct island constraints).

The goal of this paper is to re-examine English NS’ and Korean NNS’ sensitivity to English island sentences with experimental approaches. Unlike Sprouse et al. (2012) and Kim (2015) which took the Likert scale in the experiments, this study adopted magnitude estimation (ME). Since this paper adopted a different method of measurement, the statistical analyses were also different from these two previous studies.

This paper is organized as follows. In Section 2, previous studies were reviewed focused on English island constraints by English NS and Korean NNS. Section 3 introduced the experimental design and research method. Section 4 illustrated the analysis results with graphs and statistical values. Section 5 included the discussions and Section 6 summarized this paper.

Since the main focus of this paper was on the comparison of (in)sensitivity to the English island sentences by NS and Korean NNS, the current study did not contain the discussion of L1 transfer effects (here, the transfer effects by Korean) in detail. However, this kind of syntactic analysis can be the basis of the studies on the L1 transfer effects, because they might also appear in the acquisition of English island constructions by Korean NNS.

2. Previous Studies

2.1 Island Constraints in English

Canonical wh-questions in English can be constructed by moving wh-words or wh-phrases from its base position to the sentence-initial [Spec, CP] position, as shown in (1).

(1)  a. John loves whom?
    b. Whom does John love ___?

In (1b), the extracted position is marked with ___, and it is often called a trace or a gap. Though it is known that the movement operation is usually unbounded, there are also some cases where any phrase cannot cross certain kinds of syntactic boundaries. Ross (1967) collectively called these types of syntactic domains islands.1)
Since the advent of Generative Grammar, there have been a lot of proposals to explain the syntactic phenomena in the formal syntax (Chomsky 1973, Chomsky 1986, Rizzi 1990, among many others). These studies tried to unify various types of island constraints under a set of rules or generalizations. One of such ideas was the Subjacency Condition (Chomsky 1973), and it mentioned that a $wh$-phrase could not cross two or more bounding nodes in one step. Here, the bounding nodes were noun phrases (NP) and sentences (S) in English. In sum, an island could be defined as a syntactic domain that contained two or more bounding nodes.

In addition to theoretical discussions on island constraints, there have also been several experimental approaches to English island phenomena by native speakers, and Sprouse et al. (2012) was one of such studies. This study took an experimental approach to island sentences and investigated NS’ sensitivity to island constraints. The study adopted $2 \times 2$ factor combinations in (2) and tested four different types of island constraints in English, which were listed in (3)-(6).

(2) Factor Combinations in Sprouse et al. (2012)
   a. NON-ISLAND | MATRIX
   b. NON-ISLAND | EMBEDDED
   c. ISLAND | MATRIX
   d. ISLAND | EMBEDDED

(3) 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 ___?

(4) 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 ___?

(5) 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?

1) Seven types of islands were identified in Ross (1967): complex NP, adjunct islands, $wh$-islands, subject islands, left branch islands, coordinate structure islands, and non-bridge islands.
(6) 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?

Here, the wh-phrases were marked with the italic font and their gap positions with __.

A total of 173 English NS joined in the experiment, and their acceptability judgments were measured with 5-point Likert scale. The scores were transformed into z-scores and the statistical analysis results were obtained as Figure 1.

![Figure 1](image_url)

**Fig. 1.** Analysis Results in Sprouse et al. (2012)

As Figure 1 showed, the z-scores in non-island sentences were much higher than those with island sentences, and those in matrix clauses were much higher than the scores in embedded ones, except in the [non-island, embedded] of the Subject island constraints. More important finding showed that the differences between non-island and island sentences in the embedded clauses were bigger than those in the matrix clauses and that the differences among them were statistically significant ($p < .0001$).

According to the analysis in this paper, the reason for differences between two types of clauses (embedded and matrix clauses) was that the island effects were added to the differences between embedded and matrix clauses. These analysis results demonstrated that NS clearly identified all of four island constraints.
2.2 English Island Sentences by Korean NNS

Since Ross’s study on the island constraints in English (1967), there have been lots of studies on the existence of island phenomena in other languages. Accordingly, there have been lots of studies on the status of island phenomena in Korean, and there are two opposite positions in the previous approaches. Some claim that Korean clearly has island constraints (Lee 1982, Han 1992, Park 2001, Hong 2004, Park 2009). On the other hand, other scholars claim that no island effect exist in Korean (Sohn 1980, Kang 1986, Suh 1987, Chung 2005, Hwang 2007, Yoon 2011, Yoon 2012, Kim 2013). There have also been a lot of studies which examined the L1 influences on the L2 acquisition, and there have also been a few studies on the influences of the Korean language on the acquisition of the island constraints in English by the Korean NNS.

For example, Kim (2015) also conducted an experimental method to the Korean NNS. A total of 50 Korean students participated in the experiment, who resided in Seoul. Their proficiency level were identified with the TOEIC (Test Of English for International Communication) scores, and the students with more than 750 points were selected for the experiment. She employed 5-point Likert scale to measure the acceptability scores of the Korean NNS. She also contained all of four types of island sentences and analyzed the data with Generalized Linear Mixed-effects Models (GLMM). Figure 2 summarized the analysis results.

![Analysis Results in Kim (2015)](image)

Fig. 2. Analysis Results in Kim (2015)

Through the analysis, she observed that the Korean NNS clearly identified Whether islands and Subject islands but the same informants did not identify the Complex NP islands and the Adjunct islands.²)
3. Research Method

3.1 Research Questions and Hypothesis

Though Kim (2015) was meaningful in that the study tried to examine the Korean NNS’ sensitivity to English island constraints, the study did not include an experiment which measured NS’ intuitions to the island sentences. In addition, if the identical designs of experiments and the same sets of sentences had been used for both NS and NNS, it would have been possible to compare the degree of sensitivity to English island constraints by both groups of speakers. This paper tried to conduct such experiments. The identical experimental design and the same set of sentences were used for both NS and NNS, and the results was compared and statistically analyzed.

This paper has the following research questions and hypothesis.

(7) Research Questions
a. Do both groups of speakers (NS and NNS) clearly identify four types of English island sentences?
b. If not, which group of speakers did not identify which island constraints?
c. Is the sensitivity of both groups of speakers similar or not?

(8) Hypothesis
a. If both groups of speakers (NS and NNS) clearly identify four types of English island constraints, there will be statistically significant differences between island sentences and non-island sentences.
b. If not, there are some cases where the differences between island sentences and non-island sentences are not statistically significant.
c. If one group of speakers are more sensitive to English island constraints, the gap between island sentences and non-island sentences will be bigger.

To answer the research questions in (7) and to scrutinize the validity of the hypotheses in (8), two experiments were performed.

3.2 Materials

This paper followed the experimental designs in Sprouse et al. (2012). The experiment had a 2×2 design: Island constraint (Absence vs. Presence) and Location of wh-word (Matrix clause vs. Embedded clause). Though basic target sentences came from Sprouse et al. (2012), more sets of target sentences were constructed. Along with these target

---

2) Though Kim (2015) contained the plots, four plots in Figure 2 were drawn again for easy comparisons, with the means and standard deviations in Kim (2015).

3) In addition Kim (2015), B. Kim (2015; Sensitivity to Islands in Korean-English Bilinguals) conducted another types of experiments which investigated the acceptability judgments of Korean heritage speakers. This study was not mentioned in this paper, since only two type of island constraints (Wh-islands and Adjunct islands) were investigated. On the other hand, Kim (2015) contained the experimental results of all of four types of island constraints in English.
sentences, the identical number of filler sentences were made. The fillers were made up of both grammatical and ungrammatical sentences that were not related with the island constraints but included the subordinate conjunctions. A total of 160 sentences were constructed (4 island types × 4 sentence types × 5 repetitions × target/filler).

3.3 Participants

A total of 200 informants participated in the experiments. NS were the students (age: $m=20.340$, $sd=0.684$) resided in Miami, OH, USA; and NNS were the students (age: $m=21.737$, $sd=2.342$) resided in or around Daejeon area, South Korea. All of them were current university students. The TOEIC scores of Korean NNS were over 600 points, and they have studied English 9-13 years in Korea. The experiments in this study were approved by the Institutional Review Board (IRB) of the Hannam University (#17-04-01-0201). All subjects involved gave their informed written consent.

3.4 Procedures

The experiments were performed via online survey using SurveyGizmo. All the participants (both NS and NNS) were first had to fill out a simple one-page survey which contained biographical information (age and gender) and dialect(s), together with the consent form for participating in the current experiment. After they tuned their acceptability judgement using a sample set of sentences, they went into the main task. The main task was an acceptability judgment task (i.e., intuition test) using the magnitude estimation (ME) method. Among the a few different ME methods, the current study utilized a line drawing where the participants drew different lengths of lines to indicate the naturalness of a sentence.

Unlike previous studies such as Sprouse et al. (2012) and Kim (2015) which used the Likert scale, this paper adopted an ME method. Recently, many scholars have had an interest in measuring native speakers’ grammaticality/acceptability judgment on syntactic data (Bard et al., 1996, Schütze 1996, Cowart 1997, Keller 2000). The studies pointed out the problem of using the Likert scale in measuring the grammaticality/acceptability judgment as follows. First, Likert scale had limited resolution. For example, if informants might feel that the grammaticality/acceptability of a certain sentence was somewhere between two different scales (something like 2.7 or 4.5), these gradient ratings were not available in the Likert scale. Second, the method used ordinal scales, and it is not possible to guarantee that the interval between * and ** represented the same difference of impressions as that between ? and ??. Third, the Likert scale might limit our ability to compare results across the experiments. The range of grammaticality/acceptability for a set of sentences might be fitted to the scale, and what counts as ?? for a sentence might differ from what corresponds as ?? for another set of sentences.

Specifically, the acceptability scores were measured as follows. The participants were given the sentences as in Figure 3.

---

4) https://www.surveygizmo.com
A (big) dot was located at the center of the line, and the participants moved the dot towards left or right depending on the acceptability as in Figure 4.

![Fig. 4. Line Drawing (ME) Methods (After Drawing)](image)

Then, the corresponding score (here, 76) was recorded for this sentence. Following Sprouse (2008), modulus sentences were presented each page so that the participants could use them as reference sentences.

### 3.5 Statistical Analysis

After the experiments, all the data for target sentences were collected for each participant, and the data were converted into z-scores and the normality was verified for each group of sentences. When the normality tests were conducted, it was found that most of the acceptability scores did not follow the normal distribution, even after z-score transformation. Accordingly, a Generalized Linear Model (GLM) had to be used with a Gaussian distribution (a non-parametric test) in the analysis of our data. The analyses were done with R (R Core Team, 2018).

### 4. Analysis Results

#### 4.1 Overall Tendency

Before going into the results of statistical analyses using GLM, it is necessary to check overall tendencies of island sentences in both groups of speakers. Figure 5 illustrates the overall picture of island sentences by American NS, which was drawn following the tradition in Sprouse et al. (2012).

As you can see, the overall tendency of Figure 5 is similar to that of Figure 1. This implies that the sensitivity to the island constraints by American NS was similar in the experiment of Sprouse et al. (2012) and that of current study.

Figure 6 shows the overall picture of island sentences by Korean NNS. As you can see, the overall tendency of Figure 6 is different from that of Figure 2. Although the p-values of Complex NP and Adjunct islands were not statistically significant in Figure 2, the p-values of all islands were significant in Figure 6. This implies that the sensitivity to the island constraints by Korean NNS in the experiment of Kim (2015) was different from that of the current study.

Although Figure 5 and Figure 6 showed the overall tendencies, it would be desirable to examine each island constraint with more specific statistics. The following four sections provided the specific statistical values for each island constraint, which was based on the GLM analyses.

5) Someone might ask why a Generalized Linear Mixed-effects Model (GLMM) was not taken as in Kim (2015), where both speaker and item variations were considered as random factors. Actually, we also conducted a GLMM with the identical data set, but the goodness-of-fit was not much improved. Accordingly, we focused on the fixed factors, not on the random factors.
Fig. 5. Analysis Results of American NS

Fig. 6. Analysis Results of Korean NNS
4.2 Whether Islands

Table 1 and Table 2 were the GLM analysis results for Whether island constraint in Figure 5 and Figure 6 respectively.

<table>
<thead>
<tr>
<th>Table 1. GLM Analysis Results (English NS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Island</td>
</tr>
<tr>
<td>Location:Island</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. GLM Analysis Results (Korean NNS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Island</td>
</tr>
<tr>
<td>Location:Island</td>
</tr>
</tbody>
</table>

As the analysis results showed, both factors Location and Island influenced the acceptability scores, and their effects were statistically significant (p<.001). It implied that the acceptability scores for the island sentences were affected by the location of wh-extraction (Location) and by existences of island (Island). Also, note that an interaction existed between these two factors (Location:Island), and it was statistically significant (p<.001). It implied that there was an interaction between these two factors. The same pattern appeared both in Table 1 and Table 2, though the p-value of intercept was not statistically significant (p=.532) in Table 2. These results implied that both American NS and Korean NNS clearly identify the Whether island constraint.

4.3 Complex NP Islands

Table 3 and Table 4 were the GLM analysis results for Complex NP island constraint in Figure 5 and Figure 6 respectively.

<table>
<thead>
<tr>
<th>Table 3. GLM Analysis Results (English NS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Island</td>
</tr>
<tr>
<td>Location:Island</td>
</tr>
</tbody>
</table>
Table 4. GLM Analysis Results (Korean NNS)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>sd</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.015</td>
<td>0.017</td>
<td>0.893</td>
<td>.372</td>
</tr>
<tr>
<td>Location</td>
<td>-0.499</td>
<td>0.017</td>
<td>-30.429</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Island</td>
<td>-0.153</td>
<td>0.017</td>
<td>-9.325</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Location:Island</td>
<td>-0.096</td>
<td>0.017</td>
<td>-5.851</td>
<td>&lt;.001 ***</td>
</tr>
</tbody>
</table>

As the analysis results illustrated, both factors Location and Island influenced the acceptability scores, and their effects were statistically significant (p<.001). It implied that the acceptability scores for the island sentences were affected both by the location of wh-extraction (Location) and by existences of island (Island). Also, note that there was an interaction between these two factors (Location:Island), and it was statistically significant (p<.001). It implied that there was an interaction between these two factors. The same pattern appeared both in Table 3 and Table 4, though the p-value of intercept was not statistically significant (p=.374) in Table 4. These results implied that both American NS and Korean NNS clearly identify the Complex NP island constraint.

4.4 Subject Islands

Table 5 and Table 6 were the GLM analysis results for Subject island constraint in Figure 5 and Figure 6 respectively.

Table 5. GLM Analysis Results (English NS)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>sd</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.229</td>
<td>0.018</td>
<td>-12.545</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Location</td>
<td>-0.135</td>
<td>0.018</td>
<td>-7.403</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Island</td>
<td>-0.700</td>
<td>0.018</td>
<td>-38.374</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Location:Island</td>
<td>-0.222</td>
<td>0.018</td>
<td>-12.167</td>
<td>&lt;.001 ***</td>
</tr>
</tbody>
</table>

Table 6. GLM Analysis Results (Korean NNS)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>sd</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.264</td>
<td>0.017</td>
<td>-15.273</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Location</td>
<td>-0.625</td>
<td>0.017</td>
<td>-36.202</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Island</td>
<td>-0.250</td>
<td>0.017</td>
<td>-14.459</td>
<td>&lt;.001 ***</td>
</tr>
<tr>
<td>Location:Island</td>
<td>-0.166</td>
<td>0.017</td>
<td>-9.621</td>
<td>&lt;.001 ***</td>
</tr>
</tbody>
</table>

As the analysis results presented, both factors Location and Island influenced the acceptability scores, and their effects were statistically significant (p<.001). It implied that the acceptability scores for the island sentences were affected by the location of wh-extraction (Location) and by existences of island (Island). Also, note that there was an interaction between these two factors (Location:Island), and it was statistically significant (p<.001). It implied that there was an interaction between these two factors. The same pattern appeared both in Table 5 and Table 6. Here, also note that the p-value of intercept was statistically significant (p<.001) also in Table 6. These results implied that both
American NS and Korean NNS clearly identify the Subject island constraint.

4.5 Adjunct Islands

Table 7 and Table 8 were the GLM analysis results for Adjunct island constraint in Figure 5 and Figure 6 respectively.

<table>
<thead>
<tr>
<th>Table 7. GLM Analysis Results (English NS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>(Intercept)</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Island</td>
</tr>
<tr>
<td>Location:Island</td>
</tr>
</tbody>
</table>

Table 8. GLM Analysis Results (Korean NNS)

| Estimate | sd | t    | p     |
| (Intercept) | -0.169 | 0.017 | -9.680 | <.001 *** |
| Location | -0.604 | 0.017 | -34.516 | <.001 *** |
| Island | -0.286 | 0.017 | -16.347 | <.001 *** |
| Location:Island | -0.175 | 0.017 | -9.973 | <.001 *** |

As the analysis results demonstrated, both factors Location and Island influenced the acceptability scores, and their effects were statistically significant (p<.001). It implied that the acceptability scores for the island sentences were affected both by the location of wh-extraction (Location) and by existences of island (Island). Also, note that there was an interaction between these two factors (Location:Island), and it was statistically significant (p<.001). This result implied that an interaction existed between these two factors. The same pattern appeared both in Table 7 and Table 8. Here, also note that the p-value of intercept was statistically significant (p<.001) also in Table 8. These results implied that both American NS and Korean NNS clearly identify the Adjunct island constraint.

4.6 Analysis of DD-scores

Though the line plots (Figure 5 and Figure 6) and the statistical values (from Table 1 to Table 8) clearly demonstrated that both groups of speakers (American NS and Korean NNS) clearly identified all of four island constraints, the analysis results were not enough to examine how much each group of speakers felt the island sentences unacceptable. In order to measure the degree of unacceptability in island sentences, another statistical values were necessary.

Sprouse et al. (2012) adopted the differences-in-differences (DD) scores in Maxwell and Delaney (2003) and measured the strength of island sentences. The DD scores were calculated as follows.

(9) Calculation of DD Scores
a. \[ D1 = \text{NON-ISLAND|EMBEDDED} - \text{ISLAND|EMBEDDED} \]
b. $D2 = \text{NON-ISLAND}\text{MATRIX} - \text{ISLAND}\text{MATRIX}$

c. $DD = D1 - D2$

Sprouse et al. (2012) calculated the DD scores in each island constraint in Figure 1 and observed that the DD scores had the plus values in all of the island constraints. They mentioned that these plus values in the DD scores numerically showed that super-additive effects and that these effects clearly indicated the existence of island effects in the NS’ intuition.

Following the method in Sprouse et al. (2012), the DD scores of four island constraints were calculated both for Figure 5 and Figure 6, and the scores were enumerated in the following table.

<table>
<thead>
<tr>
<th></th>
<th>Whether</th>
<th>Complex NP</th>
<th>Subject</th>
<th>Adjunct</th>
</tr>
</thead>
<tbody>
<tr>
<td>English NS</td>
<td>0.616</td>
<td>0.651</td>
<td>0.888</td>
<td>0.959</td>
</tr>
<tr>
<td>Korean NNS</td>
<td>0.427</td>
<td>0.384</td>
<td>0.665</td>
<td>0.698</td>
</tr>
</tbody>
</table>

As the values indicated, all the DD scores had the plus values. This implies that both groups of speakers clearly identify four types of islands. However, note that the scores of Korean NNS were lower than those of American NS. It implied that Korean NNS were less sensitive to English island constraints than those of American NS.

5. Discussion

This paper employed the experimental designs and used the identical set of sentences to examine the sensitivity to the island sentences by both American NS and Korean NNS. This paper also adopted the ME method (specifically the line drawing) for the measurement of acceptability judgments by two groups of speakers. The collected data were statistically analyzed and the analysis results were presented with graphs and tables in Section 4.

Figure 5 and Figure 6 demonstrated that both American NS and Korean NNS clearly identify all of four types of island constraints in English. The statistical tables (from Table 1 to Table 8) numerically validate this fact. The $p$-values for two individual factors Location and Island were statistically significant ($p<.001$), and it implied that the acceptability scores of the island sentences were affected both by the location of $wh$-extraction (Location) and by existences of island (Island). Also, there was an interaction between these two factors (Location:Island), and it was statistically significant ($p<0.001$). It implied that there was an interaction between these two factors. The same tendency were observed in all types of the island constraints. These analysis results presented that both American NS and Korean NNS clearly identify all of four types of island constraints in English.

In two figures in Section 4.1, the D1 scores (NON-ISLAND|EMBEDDED $- \text{ISLAND|EMBEDDED}$) of Figure 5 looked much bigger than those of Figure 6. It might implied that American NS are more sensitive to island sentences than Korean NNS. The DD scores were calculated for all of four types of island constraints (for each group of speakers) and they were enumerated in Table 9. The analysis results implied that the scores of Korean NNS were lower than those of American NS. It implied that Korean NNS were less sensitive to English island constraints than those of American NS.
The lower DD-scores might be originated from the L1 transfer effects (Selinker 1969, Odlin 1989, Odlin 2003). However, further studies will be necessary to uncover whether the L1 transfer effects caused the lower DD-scores in Figure 6 (by Korean NNS).

On the other hand, the overall patterns of Figure 6 were different from those in Figure 2 which came from Kim (2015). Kim (2015) reported that Korean NNS were not sensitive to Complex NP island and Adjunct island, but the current study mentioned that Korean NNS clearly identify all of four types of island constraints in English. There might be various reasons. Park and Lee (2018) listed two possible reasons. The first was related with the measuring methods of acceptability scores. Kim (2015) used a 5-point Likert scales, whereas this paper adopted the ME method. As pointed out by many scholars, the Likert scale had many problems in the measurement of grammaticality/acceptability scores (Stevenson, 1975; Johnson, 2008), and some previous studies (Lodge 1981, Johnson, 2008) proposed the ME method into the study of sentence acceptability because the ME method had more fine-grained scale ranges than the Likert scale. The second reason might be originated from the statistical methods in Kim (2015). Kim (2015) seemed to treat the Likert scales as interval/ratio scales. However, as many studies mentioned, Likert scales were not treated as interval/ratio variables (which could calculate mean and standard deviation) but must be treated as ordinal variables (Argresti 2007, Argresti 2010, Gravetter and Wallnau 2013, Lee 2016). Many textbooks in statistics often said that the Likert scales should be considered as interval or ratio scales. Yes, it might be. However, in order to apply this maxim to the Likert scale, the normality assumptions must be satisfied. Depending on whether the data followed the normal distribution or not, the next statistical analysis were determined. In Kim (2015), there was no mentions about this.

Now, returning to the research questions in (7), what did the analysis results imply? The first question was whether American NS and Korean NNS clearly identify four types of English island sentences or not. The hypothesis in (8a) mentioned that if both groups of speakers distinguished four types of English island constraints, there will be statistically significant differences between island sentences and non-island sentences. The two graphs in Section 4.1 and tables from Section 4.2 to Section 4.5 showed that both groups of speakers clearly identified all of four island constraints in English. This was against Kim’s observations (2015) that Korean NNS did not identify the Complex NP islands and the Adjunct islands. All of these results supported the hypothesis in (8a), and the answer for the first research question would be yes.

The second question was which group of speakers did not identify which island constraints if either American NS or Korean NNS did not clearly identify four types of English island sentences. The hypothesis in (8b) mentioned that there are some cases where the differences between island sentences and non-island sentences are not statistically significant in those cases. Since both groups of speakers clearly identified all of four island constraints in English, the answer for the second research question would be not necessary.

The third question was if the sensitivity of both groups of speakers were identical or not. The hypothesis in (8b) mentioned that if one group of speakers is more sensitive to English island constraints, the gap between island sentences and non-island sentences will be bigger. Table 9 provided the answer. Although it was a fact that the DD-scores of Korean NNS were lower than those of American NS, the differences were not statistically significant. It implied that both groups of speakers felt similar degrees of unacceptability in island sentences though the DD-scores of Korean NNS were lower than those of American NS. All of these results did not support the hypothesis in (8c), and the answer for the third research question would be that the sensitivity of both groups of speakers was identical.
6. Conclusion

In this paper, it was examined how (American) NS and Korean NNS process English island sentences using experimental methods. This paper followed the basic design in Sprouse et al. (2012), and the same sets of sentences were used for both groups of speakers. Unlike previous studies such as Sprouse et al. (2012) and Kim (2015), this paper applied the ME method, especially a line-drawing method. After the experiments, all the data were collected for target sentences. The collected acceptability scores were transformed into z-scores, and statistical analyses were applied including the normality tests and the GLM.

Through the experiments, the followings were found: (i) Korean NNS clearly identified all of the four types of island constraints in English and (ii) DD-scores of Korean NNS were smaller than those of NS. The analysis results of Korean NNS were against Kim (2015), which mentioned that Korean NNS clearly did not identify Complex NP and Adjunct island constraints in English. The analysis results of DD-scores opened a possibility that Korean NNS were slightly less sensitive to island sentences than American NS. Further studies were required to reveal the discrepancies between American NS and Korean NNS. Although many issues still remain in the study of island phenomena in natural languages, this paper uncovers one aspect of island constraints in English, especially on the similarity and differences between American NS and Korean NNS.

References


Lee, Yong-hun, Instructor
99 Daehak-ro, Yuseong-gu, Daejeon 34134, Korea
Department of English Literature and Language
Chungnam National University
E-mail: ylee@cnu.ac.kr

Yeonkyung Park, Graduate Student
70 Hannam-ro, Daedeok-gu, Daejeon 34430
Department of English Language and Literature
Hannam University
E-mail: withbyk@gmail.com