Processing English Island Sentences by Korean EFL Learners

Yeonkyung Park
Hannam University
70 Hannamro, Daedeok-gu
Daejeon 306-791, Korea
withbyk@gmail.com

Yong-hun Lee
Chungnam National University
99 Daehak-ro, Yuseong-gu
Daejeon 305-764, Korea
yleeuiuc@hanmail.net

Abstract
This paper took an experimental approach and investigated how Korean EFL learners process the English island constructions. Since there are some controversies on the existence of the island effects in Korean, the L1 transfer effect may make it difficult for the Korean EFL learners to learn island constructions in English. To examine if the difference between English and Korean affects the acquisition of English island phenomena: Complex-NP, whether, subject, and adjunct island. The acceptability scores of Korean EFL learners were measured with Magnitude Estimation (ME). Then, the collected data were statistically analyzed. The analysis results showed that, unlike previous studies, the Korean EFL learners correctly identified all of the English island constructions. This finding showed that the island status of the Korean language did not affect the acquisition of island constructions in English.

1 Introduction
Since Ross’s identifications of island constraints in English (Ross, 1967), there have been a lot of debates on the existence of island constraints in other languages. Some languages were believed to contain some island effects, while other languages (e.g. Chinese, Korean, or Japanese) were doubtful about the existence of island effect.

The status of island effects of the L1 (the mother tongue) also may influence the acquisition of L2, since it was well-known that the knowledge of L1 might influence the acquisition of L2, which was known as the L1 transfer effects (Selinker, 1969; Odlin, 1989; 2003). Korean students learn English as Foreign Language (EFL), since English is not an official language in Korean. 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). ¹ Then, the question is whether the island status of Korean may influence the acquisition of the constructions in English. To answer this question is also crucial from the psycholinguistic point of view, since there might be different psycholinguistic or cognitive processes when people produce or understand the island constructions in their native language (L1) and another language (L2).

In order to investigate whether the L1 transfer effects also appear in the acquisition of English island constructions, an experiment was designed where the acceptability scores of the Korean EFL learners were measured with the ME method. Then, the collected data were statistically analyzed with R.

This paper is organized as follows. In Section 2, previous studies are reviewed. Section 3 includes the experimental design, research materials and research method. Section 4 enumerates the analysis

¹ Similar kinds of controversies exist also for Japanese. Nishigauchi (1990) and Watanabe (1992) claimed that there were island constraints in Japanese, but Ishihara (2002) and Sprouse et al. (2011) mentioned that this language had no island constraint.
results. Section 5 contains discussions, and Section 6 summarizes this paper.

2 Previous Studies

2.1 Island Effects in Korean

Since Ross (1967) identified the island constraints in English, there have been a lot of studies on the existence of island phenomena in other languages. These previous studies focused on examining if the island constraints existed in their languages and why the language escaped the island constraints when the language did not show the island phenomena.

Korean is no exception. There have been lots of previous studies on the island constraints also in Korean, but there are two opposite positions in the previous approaches. Some claimed that Korean has island constraints (Lee 1982; Han 1992; Hong 2004; Park, 2001, 2009). Hong (2004) proposed two diagnostics for syntactic movements: island and intervention effects. He mentioned that Korean also has an island effects. Park (2001) and Park (2009) claimed that matrix sluicing in Korean was island-sensitive, through examining the sluicing constructions in Korean.

On the other hand, other scholars claimed that there is no island effect in Korean (Sohn, 1980; Kang, 1986; Suh, 1987; Hwang, 2007; Chung, 2005; Yoon, 2011, 2012; Kim, 2013). Yoon (2011, 2012) identified two novel environments where *wh*-phrases had no island effects: the declarative intervention contexts and the embedded contexts. Kim (2013) investigated *wh*-islands in the relative clauses, and he claimed that the fact that Korean escaped the island constraint could be explained by a semantico-pragmatic constraint.

2.2 Experimental Approaches to Islands

Recently, as computer technology and statistical tools develop, many researchers had an interest in measuring native speakers’ intuition on syntactic data objectively and scientifically (Bard, Robertson, and Sorace, 1996; Schütze, 1996; Cowart, 1997; Keller, 2000). This research method was also applied into the study of island constructions, and lots of fruitful facts have been discovered through experimental approaches.

Sprouse et al. (2012), for example, adopted an experimental approach to island constructions and examined native speakers’ intuition. They adopted 2×2 factor combinations in (1) and investigated four types of island constraints using the sentences in (2)-(5) (Sprouse et al., 2012:87-8).

(1) Factor Combinations
a. NON-ISLAND | MATRIX
b. NON-ISLAND | EMBEDDED
c. ISLAND | MATRIX
d. 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 target sentences, they measured the acceptability scores of 173 native speakers. Through the experiments and their analysis, they obtained the following results (Sprouse et al. 2012:100).
These analysis results illustrated (i) that native speakers showed more acceptability for non-island structures than island structures both in matrix and embedded clauses and (ii) that the differences of acceptability scores became greater in embedded clauses rather than matrix clauses. All of these observations demonstrated that there were clearly island effects in English.

There were also some studies on the acquisition of the English island constructions by the Korean EFL learners. For example, Kim B. (2015) studied the acquisition of English island constructions by Korean-English bilinguals with an experimental approach and their statistical analysis. Sixty-three Korean-English bilinguals and sixty native speakers of English participated in the experiments. Here, bilinguals were either US-born or Korea-born who moved to the U.S. between ages 0 to 14. Based on their ages of arrival (AoA) to the U.S., bilinguals were divided into three groups: Heritage (AoA 0-5), Early (AoA 6-10), and Late (AoA 11-14). The experimental study demonstrated that all the group of speakers clearly distinguished four types of island constraints in Figure 1 (i.e., Complex NP, Whether, Subject, and Adjunct). However, the intuition of Heritage speakers were the closest to the intuitions of native speakers and the Early group was closer to natives though the group were far from the natives. The study also showed that the Late group was very far from both natives and the Heritage group. These results illustrated that, as the AoA was later, the L1 transfer effects might be stronger and the effects made it difficult for the EFL learners to learn the island constructions in the target language (here, English).

Although this study succeeded to demonstrate that the L1 transfer effects became stronger as the AoA was later, this study focused on the behaviors of the Heritage speakers. Accordingly, the study did not contain enough data which were obtained from the EFL students who resided in Korean. It is also necessary to conduct a similar experiment for the EFL students who resided in Korean.

Kim H. (2015) conducted such an experiment. In her studies, a total of fifty students participated in the experiment, who resided in Korean. Their proficiency level were classified with the TOEIC (Test Of English for International Communication), and the students with more than 750 points were included in the experiment. She adopted 5-points Likert scale to measure the acceptability scores of the Korean EFL learners. She also included four types of island constructions in Figure 1 and analyzed the data with ANalysis Of VAriance (ANOVA). Through the analysis, she found that the Korean EFL learners clearly identified the Whether island and the Subject island constraints but they did not identify the Complex NP island and the Adjunct island constraints.

Although her study was meaningful in that the experiment was conducted to the students who resided in Korean, there might be some problems which could be raised from the measurement of the acceptability scores for the Korean EFL learners. As mentioned in several previous studies (such as Bard et al., 1996; Schütze, 1996; Cowart, 1997; Keller, 2000), Likert scale has several problems compared with the ME method, to be used in the acceptability judgment tasks.² First, Likert scale 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

² 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.
method uses an ordinal scale, and there is no guarantee that the interval between * and ** (ungrammatical) represents the same difference of impressions as that between ? and ?? (between grammatical and ungrammatical). The former method, on the other hand, provides judgments on an interval scale for which averages (mean value, \(m\)) and standard deviations (\(sd\)) 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. Accordingly, another type of measuring method was necessary to solve this problem. This paper adopted the ME method to solve the problems of the Likert scale.

3 Research Method

3.1 Research Question and Hypothesis

Through the experimental study, this paper wanted to investigate if the Korean EFL learners identified four types of island constraints in Figure 1.

Our research questions are as follows.

(6) Research Questions
a. Do the Korean EFL learners clearly identify four types of island constraints in English?
b. If the answer is 'no', which island constraints in English do they clearly identify and which ones are not identified?

For these questions, we made the following hypotheses.

(7) Hypothesis
a. If there is no or little L1 transfer effect, the Korean EFL learners will clearly identify all of (four types of) the island constraints.
b. If there is a L1 transfer effect, the Korean EFL learners will not clearly identify at least one of the island constraints.

To examine these hypotheses, an experiment was designed as follows.

3.2 Materials

To closely examine the English island constraints by the Korean EFL learners, the first thing to do was to make target sentences. This paper basically followed the factor combinations in (1), following the study in Sprouse et al. (2012). Accordingly, the following two factors were used in the experiment: Island constraint (Absence vs. Presence) and Location of wh-word (Matrix clause vs. Embedded clause). Since two factors were adopted and each factor had two values, the experiment had a 2\times2 design.

First of all, basic target sentences were made with the sentences in (3) and the sentences in Pearl and Sprouse (2014), but a lexical items were slightly changed. These four sentences matched with the corresponding sentences in (3), and they contained the factor combinations in (1).

Along with these target sentences, the same number of filler sentences was made. The half of the filler sentences were constructed based on the structure of the target items. However, they were not related with the island constraints. The others were composed of the filler sentences that had no relation with the purpose of the experiment. Among them, some sentences were grammatical and others were ungrammatical. At the end, a total of 128 sentences were constructed in the experiments (4 island types\times4 sentence types\times4 repetitions).

After all the target and filler sentences were constructed, random numbers were generated with the R function (from 1 to 128; 64 target sentences and 64 fillers), and each sentence was given the generated random numbers. Then, the sentences were given to the participants after the sentences were sorted based on the random number.

3.3 Procedure

The data for a total of 20 native speakers were collected from the experiment. All the participants (\(m=23.40, sd=1.23\)) resided in and around Daejeon area, South Korea. All of them were either current university students or graduates of universities in Korea.

All the participants were first asked to fill out a simple one-page survey that contains biographical information such as age, gender, and dialect(s), together with the consent form for participating in the experiment. Then they were asked to proceed to take the main task.
The main task used in the experiment was an acceptability judgment task using Magnitude Estimation (ME; Lodge, 1981; Johnson, 2008).

There are two types of ME methods: numerical estimates and line drawing. However, as Bard et al. (1996) pointed out, the participants sometimes think of numeric estimates as academic test scores, and so they tend to limit their responses to a somewhat categorical scale, rather than using a ratio scale as intended in the magnitude estimation. Accordingly, the current study adopted a line drawing method in which the participants were asked to draw different lengths of lines to indicate the naturalness (acceptability) of a given sentence (after reading the given sentence). An acceptability judgment task (also known as native speakers’ intuition test) was 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 were required to draw a line for the given sentence, according to the degree of naturalness of the given sentence.

### 4 Statistical Analysis

#### 4.1 Normality Tests and Regression Analysis

After all the data were collected from acceptability judgment tasks, the values were extracted for target sentences. Then, the normality tests (Baayen, 2008; Gries, 2013; Lee, 2016) were performed to check whether parametric tests were 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 were performed, it was found that all the data sets did not follow the normal distribution. Some were positively skewed, and other sets had a slightly bimodal distribution. Consequently, non-parametric tests had to be used in the analysis of our data.

After the normality tests were performed, a (generalized) regression test (GLM) was performed. According to Agresti (2007), a generalized regression test is available when the distribution does not follow the normal distribution. Thus, the test was adopted to examine how each factor affects the acceptability of the sentences.

#### 4.2 Complex NP Islands

Table 1 illustrated the analysis results of the GLM analysis.

| Estimate (Intercept) | sd  | t      | p     |
|----------------------|-----|--------|-------|
| 145.1844             | 2.9011 | 50.045 | <<<.001 |
| CLAUSE -0.1406       | 2.9011 | -0.048 | 0.9614 |
| ISLAND 17.3594       | 2.9011 | 5.984  | <<<.001 |
| CLAUSE-ISLAND 5.4719 | 2.9011 | 1.886  | 0.0602 |

Table 1: GLM Analysis Results for Complex NP

As you can see in this table, the factor CLAUSE was not significant ($p=.9614$), but the factor ISLAND was highly significant ($p<.001$). The interaction between these two factors was marginally significant ($p=.0602$).

Figure 2 showed us an effect plot for this island constraint.

![Figure 2: Interaction Plot for Complex NP](image)

As you can see in this interaction plot, the overall acceptability scores became lower when the island constraint existed (i.e., Presence). The difference in the acceptability scores was bigger in the Embedded clause than in the Matrix clause. It implies that the Differences-in-Differences (DD) scores may have the plus values and that the Korean EFL learners surely identify the Complex NP island constraints in English.

#### 4.3 Whether Islands

Table 2 illustrated the analysis results of the GLM analysis.
As you can see in this table, both factors CLAUSE and ISLAND were significant ($p=0.00698$ and $p<0.001$ respectively). The interaction between these two factors was marginally significant ($p=0.05239$).

Figure 3 showed us an effect plot for this island constraint.

As you can see in this interaction plot, the overall acceptability scores became lower when the island constraint existed (i.e., Presence). The difference in the acceptability scores was bigger in the Embedded clause than in the Matrix clause. It implies that the DD scores may have the plus values and that the Korean EFL learners surely identify the Whether NP island constraints in English.

### 4.4 Subject Islands

Table 3 illustrated the analysis results of the GLM analysis.

As you can see in this table, the factor CLAUSE was not significant ($p=0.256$), but the factor ISLAND was highly significant ($p<0.001$). The interaction was not significant ($p=0.333$).

Figure 4 showed us an effect plot for this island constraint.

As you can see in this interaction plot, the overall acceptability scores became lower when the island constraint existed (i.e., Presence). The difference in the acceptability scores was bigger in the Embedded clause than in the Matrix clause. It implies that the DD scores may have the plus values and that the Korean EFL learners surely identify the Subject island constraints in English.

### 4.5 Adjunct Islands

Table 5 illustrated the analysis results of the GLM analysis.

As you can see in this table, the factor CLAUSE was not significant ($p=0.256$), but the factor ISLAND was highly significant ($p<0.001$). The interaction was not significant ($p=0.333$).

Figure 6 showed us an effect plot for this island constraint.

As you can see in this interaction plot, the overall acceptability scores became lower when the island constraint existed (i.e., Presence). The difference in the acceptability scores was bigger in the Embedded clause than in the Matrix clause. It implies that the DD scores may have the plus values and that the Korean EFL learners surely identify the Adjunct island constraints in English.
constraint existed (i.e., Presence). The difference in the acceptability scores was bigger in the Embedded clause than in the Matrix clause. It implies that the DD scores may have the plus values and that the Korean EFL learners surely identify the Adjunct island constraints in English.

5 Discussion

The analysis results in Section 4.2-4.5 illustrated different aspects that Kim H. (2015) observed in her experimental studies. In her study, she found that the Korean EFL learners clearly identified the Whether island and the Subject island constraints but they did not identify the Complex NP island and the Adjunct island constraints. However, in this study, the Korean EFL learners clearly identified all of the island constraints.

Then, where did the differences come from? There may be two types of sources which made the differences. The first one might come from the methods of measuring the acceptability scores. Kim H. (2015) used a 5-point Likert scales, while this paper adopted the ME method. Although the Likert scales were widely used in previous studies, they had some shortcomings as mentioned in Section 2.2. Even though we did not take the problems into consideration, the ME method had more fine-grained scales than the Likert scale. Accordingly, more fine-grained differences in the acceptability scores were represented in the ME method, whereas the differences might be lessened or neutralized in the Likert scale, especially in the Complex NP and the Adjunct island constraints. The second origin came from the statistical method. In Kim H. (2015), the collected data were analyzed with z-transformation. Originally, the Likert scale was an ordinal variable (Lee, 2016). Consequently non-parametric tests had to be applied. In order to solve the problem, Kim H. (2015) employed a z-transformation, which made the ordinal variables like the ratio variables. However, z-transformation was also a transformation. That is, the data might be distorted during the transformation processes. This paper, on the other hand, did not apply any kind of transformation to the collected data. Since the acceptability scores were ratio variables (Lee, 2016), the normality tests were applied. Since the result was that the distributions did not follow the normal distributions, GLM methods were applied. Therefore, no transformation was adopted here, and the data were not distorted. Accordingly, the analysis results in this paper could be said to be more accurate than those in Kim H. (2015).

Now, let’s see what answers can be provided to the research questions in (6) and Hypothesis in (7) along with the analysis results.

For the first question, the analysis results said that the Korean EFL learners clearly identified four island constraints in English, which was different from the analysis results in Kim H. (2015). For two hypotheses in (7), it could be said that there was no or little L1 transfer effect, since the Korean EFL learners clearly identified four island constraints in English. This implies that the unstable status of island constructions in Korean did not affect the acquisition of island construction in English.

6 Conclusion

In this paper, it was closely examined how the Korean EFL learners identified the English island constructions. Four types of island constructions (Complex NP, Whether, Subject, and Adjunct) were taken, and two linguistic factors (CLAUSE and ISLAND) were taken in the analysis, which made the experiment have a 2×2 design.

Based on this design, an acceptability judgment task was performed, where the data for 20 Korean native participants were collected with the ME method. After the experiments, all the values were extracted for target sentences and they were analyzed with R.

Through the experiments, it was found that the Korean EFL learners correctly identified all of the English island constructions. This finding showed that the island status of the Korean language did not affect the acquisition of island constructions in English.

References

Akira Watanabe. 1992. Wh-in-situ, Subjacency, and Chain Formation, MIT Occasional Papers in Linguistics 2. MIT Press, Cambridge, MA.

Alan Agresti. 2007. An Introduction to Categorical Data Analysis. John Wiley & Sons, Hoboken, NJ.

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

Boyoung Kim. 2015. Sensitivity to Islands in Korean-English Bilingualism. Doctoral dissertation. University of California at San Diego.
