L2 Processing, Proficiency, and Acceptability Judgment of *Wh*-Island Sentences

Hye-ryeong Hahn*

The current study explores the effects of processing demands and proficiency on second language (L2) learners' acceptability judgment of wh-island sentences. A total of 65 adult Korean learners of English and ten native speakers (NSs) of English participated in an experiment that combined self-paced reading and acceptability judgment. They were presented with *wh*-island and non-island sentences that varied by whether the intervening nominal constituents inside the dependency were lexical noun phrases (NPs) or pronouns. The results showed that high-proficiency learners gave a higher rating to non-island sentences than to *wh*-island sentences, which is comparable to the NSs' rating pattern, whereas low-proficiency learners showed a chance-level performance. The learners preferred the lexical condition to the pronoun condition, whereas the NSs exhibited the opposite preference. In the reading time (RT) analyses, both NSs and L2 learners slowed down in the lexical condition at regions where intervening NPs occurred, indicating that lexical nouns incur higher processing costs compared with pronouns. The NSs slowed down at the embedded verb region in the pronoun condition, whereas L2 learners showed no RT difference between conditions. The findings suggest that L2 learners' acceptability judgment is affected by their proficiency, but not as much by processing costs in terms of RT.

Keywords: L2 processing, proficiency, acceptability judgment, intervening material, self-paced reading

1 Introduction

Acceptability judgment has been one of the most widely-used research methods in first language (L1) and second language (L2) studies. Given that language is rule governed, its users are expected to accept sentences that conform to the rules of language and reject those that violate them.

However, there are some grey zones where language users' judgments are not categorically determined. That is, some violations of rules seem to result in clearly unacceptable sentences whereas other violations seem to result in largely undesirable but marginally acceptable sentences. Island constructions in English are prime examples. In English, an element cannot be

---

* **Hye-ryeong Hahn**, Professor, Department of English Education, Seowon University

Hye-ryeong Hahn

extracted from an island such as a subject clause island, an adjunct clause island, or a *wh*-clause island. But some island violations have a smaller effect size than other island violations. For instance, native English speakers are found to judge extractions from an adjunct island as in (1) as a more serious violation than extractions from a *wh*-island as in (2). Thus the former type of island violation is called a strong violation and the latter type a weak violation (Kluender, 1998; Szabolsci & Lohndal, 2017; Szabolsci & zwart, 1993; Tabor et al., 2020).

(1) *Which column did the reporter read [because the editor wrote __]?
(2) ?Which column did the reporter wonder [whether the editor wrote __]?

More interestingly, some *wh*-island violations are judged as more acceptable than other *wh*-island violations (Belletti & Rizzi, 2013; Hofmeister & Sag, 2010; Kluender, 1998; Kluender & Kutas, 1993; Pesetsky, 1987; Rizzi, 1990). For example, NSs are found to rate (3) as more acceptable than (4) although both violate the *wh*-island constraint.

(3) ?Which patient did the nurse ask [whether the doctor treated __]?
(4) ??Who did the nurse ask [whether the doctor treated __]?

Such gradience in island effects suggests that acceptability judgment is affected by other factors than syntax per se. One of the most promising accounts for the graded acceptability so far is a processing-based account (Hofmeister & Sag, 2010; Hofmeister et al., 2013; Tabor et al., 2020). According to processing-based approaches, acceptability-or grammaticality-of complex sentences is attributable to processing factors such as memory and processing costs. Based on L1 processing experiments, they propose that unacceptability of some complex sentences is due to excessive memory load required to process them.

If NSs' acceptability judgment is affected by processing difficulty, and if processing difficulty is closely related to memory resources, L2 learners' acceptability judgment is even more vulnerable to processing factors, as they suffer from more serious working memory shortage (McDonald, 2006; VanPatten, 2004). Further, L2 judgments might vary between learners with high memory span and those with low memory span. Judgments are also likely to vary between proficient learners and less proficient learners: complex sentences will be more acceptable for proficient learners whose L2 processing has become more automatic compared to less proficient learners.

Despite numerous studies on L2 acceptability judgment and recent increase in studies on L2 processing, the relationship between L2 processing and L2 acceptability is still an understudied research area. The current paper explores whether and to what extent L2 acceptability judgment is interrelated with processing difficulty and proficiency. Specifically, it addresses whether

Korean EFL learners' acceptability judgments of *wh*-island constructions are affected by (i) the amount of processing demands and (ii) L2 proficiency. To this end, the study combined two experiments—acceptability judgment and self-paced reading—and investigated how learners at different proficiency levels process and judge *wh*-island sentences with different processing demands.

2 Background

2.1 *Wh*-island sentences and processing difficulty

Wh-questions in English involve movement of a *wh*-phrase from its argument position, where it is base-generated, to a clause-initial position as in (5) and (6).

(5) Who did Bill meet __ yesterday?
(6) I want to know who Bill met __ yesterday.

A *wh*-phrase can also move across a clausal boundary as in (7), but this kind of long-distance movement is not allowed when the embedded clause is headed by another *wh*-phrase as in (8), leading to *wh*-island effects (Ross, 1987).

(7) Who did Tom say that Bill met __ yesterday?
(8) *Who did Tom ask whether Bill met __ yesterday?

In generative linguistics, this and other island constraints on movement are generalizable in terms of Subjacency. According to the Subjacency principle, an element cannot be extracted across more than one bounding nodes, which are IPs and CPs in English. In (9), *who*, which is base-generated in the embedded clause moves to the sentence-initial position via the intermediate [Spec, CP] as a landing site. Each movement involves crossing a single bounding note and thus does not violate Subjacency. By contrast, the embedded *wh*-phrase in (10) cannot use the intermediate [Spec, CP] as a landing site because it is already occupied by another *wh*-element *whether*. As *who* has to cross two IPs in a single movement, the movement violates Subjacency, and therefore is considered unacceptable.

(9) Who did [IP Tom say [CP___ that [IP Bill met__ yesterday]?
(10) *Who did [IP Tom say [CP whether [IP Bill met __ yesterday]?

The rule-based account for *wh*-island effects have been contended by other researchers who noted that the acceptability of *wh*-island sentences are gradient rather than categorical. For example, *wh*-island effects are weaker in

(11) than in (12), although both violate Subjacency (Hofmeister & Sag, 2010; Kluender, 1998; Kluender & Kutas, 1993; Pesetsky, 1987; Rizzi, 1990).

(11) ?Which course did the girl ask whether she could take?
(12) ??What did the girl ask whether she could take?

The varying acceptability of *wh*-island sentences cannot be properly explained by a rule-based account due to its categorical nature. A competing account, the processing-based account, attempts to explain the variance in the acceptability of *wh*-island sentences as well as their general unacceptability. According to processing-based accounts, *wh*-island sentences are generally unacceptable because they are extremely difficult to process. To begin with, processing them involves a long-distance dependency resolution between the fronted *wh*-element (i.e., the filler) and its argument position from which it has been moved (i.e., the gap). Put differently, the extracted *wh*-phrase must be integrated with its predicate in the embedded clause to arrive at a successful interpretation. In (11), for example, in order for the *wh*-filler "which course" to be interpreted, it must be linked with its base-generated position and be associated with the verb "take." To do this, the parser must hold the *wh*-phrase in memory and process the intervening elements until it encounters its predicate, which increases processing costs. While processing, the parser also has to cross a clausal boundary, which is known to be costly (Frazier & Clifton 1989; Ginzburg & Sag 2000; Kluender, 1998). Even worse, it encounters another *wh*-phrase that intervenes between the fronted *wh*-phrase and its embedded verb, which is highly distracting as there exist two potential targets of the dependency (Belletti & Rizzi, 2013; Gordon et al., 2001; Rizzi, 2001).

Due to the excessive disruption while processing the constituents within the dependency, and due to the limited working memory capacity, the *wh*-element deteriorates as the parser progresses towards the gap site (Grodner & Gibson 2005; O'Grady 2008; Roberts & Gibson, 2002). The memory shortage makes it difficult for the *wh*-argument to be interpreted in association with its predicate, leading to processing failure. The general unprocessibility of *wh*-island sentences, according to processing-based accounts, is what causes their general unacceptability.

Still, the processing difficulty involved with *wh*-island sentences can be alleviated depending on the context. For example, island sentences containing *which-N* as in (11) may be more processible than those containing a *wh*-pronoun as in (12) (Belletti & Rizzi, 2013; Hofmeister & Sag, 2010; Kluender, 1998; Kluender & Kutas, 1993; Pesetsky, 1987; Rizzi, 1990). Some researchers claimed that *which-N* has more possibility of reactivation at the gap site than a *wh*-pronoun because the initial deep processing of the rich lexical content of *which-N* makes it easier to be reactivated at the gap site (Hofmeister & Sag, 2010; Kluender, 1998; Kluender & Kutas, 1993).

Hofmeister and Sag (2010) tested whether NSs' acceptability judgments for *wh*-island sentences are related with their processing difficulties. In their self-paced reading experiments, they presented *wh*-island sentences in two conditions so that the fronted *wh*-phrase was either a *wh*-pronoun *who* or a *which-N* sequence (e.g., *which employee*). NS participants read the sentences on a self-paced reading test and then rated the acceptability of those sentences. Their reading time (RT) analysis revealed that the RT around the embedded verb region was significantly shorter in the *which-N* condition than in the *wh*-pronoun condition, suggesting that the *which-N* condition required less processing costs. They also found that the same NSs judged sentences in the *which-N* condition as more acceptable than those in the *wh*-pronoun condition. Based on the parallelism between their acceptability ratings and RTs, they claimed that the varying acceptability of *wh*-island sentences is attributable to their varying degree of processing demands.

Processing burdens involved with *wh*-island sentences can also decrease by reducing the distance of the filler-gap dependency, as the very existence of intervening material inside the dependency creates extra processing load (Gibson, 1998, 2000; Grodner & Gibson, 2005; Just & Carpenter, 1992; Kaan, 2002; Pearlmutter, 2000). Reducing the number of intervening words or syllables is therefore one way to reduce the filler-gap distance and to alleviate cognitive demands.

The properties of the intervening material inside the dependency is another factor that can affect processing demands. Gibson and colleagues suggest that processing new discourse referents such as lexical noun phrases (NPs) that are introduced between an argument and its predicate increases the integration cost whereas pronouns do not (Gibson, 1998; Grodner & Gibson, 2005; Warren & Gibson, 2002).

Taken together, the processing-based approach predicts that the acceptability of *wh*-island sentences will increase or decrease depending on processing-related variables such as the properties of the fronted *wh*-phrase, the length/amount of the intervening material, and the properties of the intervening material.

2.2 L2 proficiency and sentence processing

Given that *wh*-island sentences are difficult for NSs to process, they can be more difficult for L2 learners because L2 learners' working memory capacity is more limited (McDonald, 2006; VanPatten, 2004). According to McDonald (2006), L2 learners experience "processing difficulties due to (1) low L2 working memory capacity, (2) poor L2 decoding, and/or (3) inadequate L2 processing speed" (McDonald, 2006, p. 381). He further argued that late L2 learners' poor acceptability judgments can be attributed to their poor processing ability.

As learners advance in their L2 acquisition, however, their L2 processing will gradually change from a controlled mode to a more automatic mode (McLaughlin et al., 1983). Their knowledge of L2 will be gradually proceduralized (Ullman, 2001). The processing costs for decoding L2 words and computing structures will gradually reduce. Processing speed will grow faster. Accordingly, more cognitive resources will be available for complex structural computations such as long-distance filler-gap dependency resolution. The improved processability then might lead them to judge *wh*-island sentences as more acceptable. It is therefore predicted that L2 proficiency will affect their processing ability, and eventually, their acceptability judgment.

2.3 Previous studies on L2 island effects

Early studies on L2 island effects have focused on whether L2 learners get access to island constraints, or subjacency (Bley-Vroman, 1989; Johnson & Newport, 1991; Lakshmanan et al., 2009; Ojima, 2005; Schatchter, 1989; White, 1992). Acceptability judgment tests were by far the most frequently used research instrument. Researchers were particularly interested in learners of English whose L1 does not instantiate Subjacency, such as Korean and Japanese. In these languages, long-distance *wh*-movement across the embedded clause is via optional movement called scrambling, not via feature-driven movement as in English. A *wh*-element originating from an embedded clause typically remains in situ in Korean. As such, the question whether rules like Subjacency, which operates on long-distance movement, are also operative in these learners' L2 system was considered an important question in SLA. Numerous studies addressed this question employing various types of acceptability judgment. Some researchers argued that Subjacency may operate for learners whose L1 instantiates it, but not for those whose L1 does not instantiate it (Bley-Vroman, 1989; Johnson & Newport, 1991; Schatchter, 1989). Others found that learners whose L1 does not instantiate Subjacency differentiated *wh*-island sentences from non-island sentences, which suggests that L2 learners also have access to Subjacency (Lakshmanan et al., 2009; Ojima, 2005; White, 1992).

While the controversy remains unsettled, the mixed findings in the literature seems to be at least partially attributable to the different levels of learners investigated in these studies. Hahn (2017) divided her Korean EFL learners into three proficiency groups and investigated whether learners' proficiency can affect the availability of the *wh*-island constraint. In her acceptability judgment experiment, she found that the high- and mid-proficiency groups gave significantly lower ratings to *wh*-island sentences than to non-island sentences. However, only the high-proficiency learners were found to correctly reject *wh*-island sentences. The low-proficiency group failed to distinguish between *wh*-island and non-island constructions, showing only

a chance-level performance on their judgments of both *wh*-island and non-island sentences.

More recent studies on L2 learners' access to island constraints have attempted to explore whether learners also process island sentences similarly to NSs (Aldosari et al. 2022; Felser et al., 2012; Johnson et al., 2016; Omaki & Schulz, 2011). Felser et al. (2012) employed eye-tracking techniques to investigate the timing pattern in L1 and L2 *wh*-island processing. They found that both NSs and proficient German learners of English showed sensitivity to extraction from island violation while processing *wh*-island sentences. However, the two groups showed different timing patterns: while the NSs were sensitive to the structural cue of the filled gap, the learners showed sensitivity to semantic features. The results led them to conclude that L2 processing is semantically mediated whereas L1 processing is syntactically mediated. Their findings highlight the contrast between L1 and L2 processing, similar to those by Clahsen and Felser (2006), who suggested that L2 learners employ a thematically-driven shallow parsing strategy whereas L1 learners use structure-based parsing. The ideas of semantically-driven L2 parsing and shallow structure were challenged by researchers who argued that L1 and L2 processing are both structure-based (Aldosari et al. 2022; Johnson et al., 2016; Omaki & Schulz, 2011). Using a self-paced reading experiment, Johnson et al. (2016) compared Korean L2 learners' time course of *wh*-island sentence processing with that of NSs. They found that Korean learners of English exhibited a time course qualitatively similar to NSs', concluding that learners can also use syntactic information similar to NSs.

While L2 studies on wh-island sentences so far have attempted to explore whether island effects are a part of their internal L2 knowledge, studies that directly delved into the effect of processing difficulty on L2 acceptability judgment are extremely rare. Hahn (2017) is one of the rare cases. She explored the relationship between L2 acceptability and processing difficulty in her acceptability judgment study. She manipulated the processing load of *wh*-sentences by varying the properties of the fronted *wh*-island (*which-N* vs. *what*) and intervening constituents (lexical NPs vs. pronouns). She found that native English speakers judged island sentences containing *which-N* as significantly better than those containing *what*, replicating Hofmeister and Sag (2010). She further found that NSs judged sentences containing pronouns as significantly better than those containing lexical nouns, which suggests that sentences with higher processing costs lowered their acceptability. On the other hand, her Korean participants did not show clear preferences except for the high-proficiency group. Only the high-proficiency learners' judgments were found to be affected by the two processing factors to some extent. Similarly to NSs, they showed preference for the *which-N* condition to the *what* condition. However, they judged *wh*-island sentences with intervening lexical NPs as more acceptable than those with intervening pronouns, exhibiting the opposite tendency in comparison to NSs. Hahn suggested that the high-proficiency

learners' preference for the lexical condition might be attributable to their lexical/semantic parsing strategy. While NSs prefer *wh*-island constructions that are less costly in terms of cognitive resources, advanced learners, she proposed, resort to semantic clues from rich lexical material rather than resources-saving pronouns in order to associate a displaced argument with its predicate.

Hahn's (2017) study, however, has some methodological limitations in that her findings were based on a single means of acceptability judgment. As she admits, the study did not actually measure the processing difficulty experienced while parsing the sentences. The processing difficulty was simply manipulated by varying conditions of the test sentences, and these different conditions were assumed to cause different amount of processing costs. However, the difference in acceptability judgment between the two conditions might have resulted from other factors than the difference in processing conditions. For example, some learners' preference for the lexical condition might be more related to their difficulties processing pronouns rather than their reliance on lexical/semantic clues. To figure out the relationship between processibility and acceptability, therefore, one needs to measure the actual processing costs by employing more rigorous psycholinguistic tools that reveal processing demands along the time course and compare the results with acceptability judgment results.

The current study aimed to address this problem and investigate whether sentences that require greater processing demands in terms of the participants' RTs are judged as less acceptable. To this end, it employed self-paced reading as a tool to reveal processing costs along the time course, in addition to acceptability judgment. The research questions are:

1) Are Korean learners of English sensitive to wh-island violation?
2) Are some wh-island sentences more acceptable to them?
3) Is their acceptability judgment of wh-island sentences affected by the amount of processing demands?
4) Are their acceptability judgment and processing of wh-island sentences affected by their proficiency level?

3 Method

3.1 Participants

A total of 65 adult learners of English and ten native English speakers participated in an experiment which combined self-pace reading and end-of-the-sentence acceptability judgment. The Korean learners were English-major undergraduates attending a local university in South Korea. Their proficiency levels varied, with their TOEIC scores ranging from 600 to 958 (*M*=770.8).

Based on their scores, they were presumed to be approximately at an intermediate or advanced level, and this proficiency range was considered to be appropriate for the current experiment, as tasks that require processing complex sentences involving long-distance wh-movement would not be appropriate for beginning level learners. The learners were later divided into two groups based on their performance on controlled items which tested whether they correctly accept grammatical, non-island long-distance *wh*-questions (see 4.1.2 for details).

The NS participants were Americans and Canadians. Half of them were English instructors at the same university, and the other half were recruited with the help of the instructors. All of them were college graduates in their 30s or 40s. They participated as a control group for baseline data as to the performance on acceptability judgment and self-paced reading. L2 participants received a gift card and NSs received cash for their participation.

3.2 Materials

The test material used in the experiment contained 18 test items and 42 filler items. The test items included 15 experimental items and three control items. The experimental items were English *wh*-island sentences containing a fronted *wh*-phrase extracted from an embedded *wh*-island clause. As in Hofmeister and Sag (2010) and Hahn (2017), all the *wh*-island clauses used in the experiment were *whether*-clauses. The *wh*-island sentences had the following sequence:

(13) Which-N did NP_1 V-ed whether NP_2 would V this morning?[1]

Sentences were presented in two conditions by manipulating the properties of the intervening material between the *wh*-filler and its gap. In the lexical condition, the two intervening NPs (one in the matrix subject position and the other in the embedded subject position) consisted of lexical nouns as in (14); in the pronoun condition, the intervening NPs consisted of pronouns as in (15).

(14) Which meeting did <u>the secretary</u> wonder whether <u>the boss</u> would cancel this morning?
(15) Which meeting did <u>she</u> wonder whether <u>he</u> would cancel this morning?

As lexical NPs are expected to require more processing costs than pronouns according to psycholinguistic findings, it was predicted that processing sentences in the lexical condition would charge more memory resources than the pronoun condition.

---

[1] The two final words were time adjuncts like *this morning* or *next week*.

Each sentence was divided into ten regions as shown in Table 1 and was horizontally presented in a single line region-by-region on the computer screen.

Table 1. Regions in the Lexical and Pronouns Conditions

	Region									
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Lexical condition	Which meeting	did	the secretary	wonder	whether	the boss	would	cancel	this	morning?
Pronoun condition	Which meeting	did	she	wonder	whether	he	would	cancel	this	morning?

In addition to the *wh*-island sentences, non-island sentences containing *that* instead of *whether* were constructed as control items, as in (16) and (17), to see if the grammatical long-distance movement across *that* is operative in their L2 system.

(16) Which meeting did <u>the secretary</u> think that <u>the boss</u> would cancel this morning?

(17) Which meeting did <u>she</u> think that <u>he</u> would cancel this morning?

A Latin square design created two experimental lists so that each participant might encounter one item in only one of the two conditions. The eighteen test sentences were interspersed with the filler items and were randomized for each trial.

3.3 Procedure

The experiment was conducted online by running the Ibex software. Participants were informed to individually join the experiment by clicking the link to the test. Each sentence was presented on the computer screen region by region, in a moving-window non-cumulative reading mode. The participants pressed the space bar to make the first region appear and proceeded to the next at their own pace. Upon reading the last region of a sentence (R10), they pressed the space bar to move to an acceptability judgement task, which asked them to rate the acceptability of the sentence just read on a 5-point Likert scale (1=*bad*, 5=*good*). Participants took 15-20 minutes to complete the entire session. Their acceptability judgment scores and RTs at individual regions were automatically recorded.

3.4 Data analyses

Participants' acceptability judgment and RT data were entered into a series of *t*-tests and ANOVAs. To determine whether L2 learners are as sensitive to *wh*-violation as NSs are, the two groups' acceptability ratings for *wh*-island

sentences and non-island control sentences were compared running *t*-tests. Each group's ratings were then compared between the lexical condition and the pronoun condition.

The L2 participants were further divided into two proficiency groups to see if learners' ratings varied depending on whether they had acquired long-distance movement from non-island embedded clauses (see 4.1.2 for details). The two groups' performances on the acceptability judgment test were compared with each other and with the NSs' by running the same *t*-test procedure. To further verify whether the effect of the intervening material varied depending on their proficiency levels, a mixed ANOVA was performed with intervening material and level as variables.

Participants' RTs were also compared across groups and conditions, focusing on (i) R3, where the first intervening NP (i.e., the matrix subject) varies between a lexical noun and a pronoun, (ii) R6, where the second intervening NP (i.e., the embedded subject) varies between a lexical noun and a pronoun, (iii) R8, the critical region, which involves the association of the fronted *wh*-element and its predicate, and (iv) R9, where a spillover effect from R8 might be observed.

Running a series of *t*-tests, for each participant group, RTs in individual regions were compared between the two conditions. NSs' and L2 learners' RTs in the lexical and pronoun conditions were compared running a mixed ANOVA with intervening material and group as main variables. To measure the effect of learner proficiency on RTs in the two conditions, a mixed ANOVA was performed with intervening material and level as main variables. Before the analyses, the RT data were trimmed by removing RTs beyond 2 standard deviations in each region, to avoid possible distortion of the whole picture due to exceptionally long RTs caused by participants' temporary distraction.

4. Results and Discussion

4.1 Acceptability judgment

4.1.1 NSs
NSs showed overall acceptance of the non-island constructions containing *that* (*M*=4.2, *SD*=1.0). They provided a higher rating in the pronoun condition (*M*=4.4, *SD*=0.9) than in the lexical condition (*M*=3.8, *SD*=1.5), but the difference was not statistically significant (*t*=1.26, *p*=.24), probably because of the small number of participants (*n*=10).

For the corresponding *wh*-island constructions, NSs provided much lower ratings (*M*=3.2, *SD*=0.8). The rating difference between island constructions and non-island constructions was significant (*t*=3.12, *p*<.05),

which demonstrates that they distinguish island-violating sentences from non-violating sentences.

As in the non-island sentences, *wh*-island sentences in the pronoun condition received a higher rating (M=3.5, SD=1.3) than those in the lexical condition (M=2.9, SD=1.2), but the difference was not significant (t=1.78, p=.11), which might also be attributable to the small number of the NS subjects. Table 2 presents NSs' acceptability ratings in the two conditions for the island and non-island constructions.

Table 2. NS Acceptability Judgment Ratings by Condition

	Lexical		Pronoun			
	M	(*SD*)	*M*	(*SD*)	*t*	*p*
Non-island	3.8	(1.5)	4.4	(0.9)	-1.26	.24
Wh-island	2.9	(1.2)	3.5	(1.3)	-1.78	.11

4.1.2 L2 learners

The L2 participants showed only a moderate acceptance toward the non-island constructions (M=3.6, SD=0.8). Their rating for non-island sentences was significantly lower than NSs' (t=2.20, p<.05). The rating difference between the lexical condition (M=3.6, SD=1.1) and the pronoun condition (M=3.7, SD=1.2) was minimal and was not statistically significant (t=0.42, p=.68). The findings seem to suggest that the learner group as a whole has not internalized long-distance *wh*-movement in their interlanguage system.

However, as the learner group was considered a mixed-ability group as mentioned, they were further divided into two groups: a high-proficiency group who had acquired long-distance *wh*-movement (Level 2) and a low-proficiency group who had not acquired it (Level 1). Acquisition of long-distance movement was operationally defined as accepting (i.e., giving a score of 4 or 5 on the 5-point Likert scale) at least two out of the three non-island control items with a *that*-complementizer. Following this criterion, 33 learners were assigned to Level 1, and 32 were assigned to Level 2.

Comparing the two proficiency groups' ratings, Level 1 did not clearly accept or reject the grammatical non-island construction (M=2.9, SD=1.1) whereas Level 2 exhibited a quite strong acceptance for the same construction (M=4.4, SD=0.6). There was a highly significant difference in judgment between the two groups (t=9.69, p<.001), which suggests that they are truly two separate proficiency groups in terms of long-distance movement.

Table 3 presents the two proficiency groups' ratings for non-island constructions in the lexical and pronoun conditions. As shown in Table 3, the difference between the lexical and pronoun conditions was not significant either in Level 1 and Level 2. Level 1 performed only at a chance level in both the lexical condition (M=2.9, SD=1.0) and the pronoun condition (M=2.9, SD=1.2); Level 2 gave high ratings both to the lexical condition (M=4.3, SD=0.6) and the pronoun condition (M=4.4, SD=0.6). The difference between conditions was not significant either in Level 1 (t=0.09, p=.93) or in Level 2 (t=0.78, p=.44).

Table 3. L2 Ratings for Non-Island Constructions by Condition

	Lexical		Pronoun			
	M	(*SD*)	*M*	(*SD*)	*t*	*p*
Level 1 (n=33)	2.9	(1.0)	2.9	(1.2)	0.09	.93
Level 2 (n=32)	4.3	(0.6)	4.4	(0.6)	0.78	.44

A mixed ANOVA confirmed that there was a highly significant main effect of level (*F*=92.59, *p*<.001), but neither the effect of condition (*F*=0.26, *p*=.61) nor interaction effect (*F*=0.10, *p*=.80) was significant.

Turning to the *wh*-island construction, L2 learners as a whole group gave it a moderately lower rating (*M*=3.4, *SD*=0.7) compared to the non-island construction (*M*=3.6, *SD*=0.7). Despite the small numerical size of the difference, their ratings for non-island sentences were found to be significantly higher than those for *wh*-island sentences (*t*=2.97, *p*<.01). The findings suggest that the learner group also distinguished the two types of long-distance *wh*-movement to some extent.

Level 2 showed a higher level of acceptance of *wh*-island sentences (*M*=3.6, *SD*=0.7) than Level 1 (*M*=3.1, *SD*=0.5). The difference in ratings was significant between groups (*t*=3.27, *p*<.01), which suggests that those who accept long-distance *wh*-movement over a non-island clause also tend to be more receptive to long-distance movement over a *wh*-island clause.

When their ratings for the *wh*-island construction and the non-island construction were compared by level, it was found that only Level 2 distinguished the island and non-island constructions. Level 2 judged the *wh*-island construction as significantly degraded than the non-island construction (*t*=4.66, *p*<.001), whereas Level 1 did not show a rating difference between the two constructions (*t*=0.65, *p*=.52). Figure 1 illustrates the two level groups' different judgment patterns for *wh*-island sentences and non-island sentences.
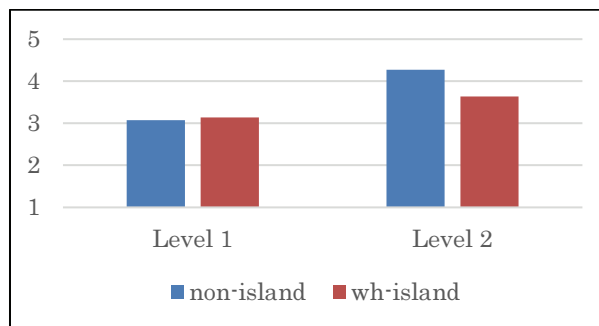

Figure 1. Acceptability ratings by level: *Wh*-island vs. non-island

The findings indicate that it was Level 2 who contributed to the whole group's differentiation between the island and non-island constructions. A mixed ANOVA with level and complementizer (i.e., *that* vs. *whether*) as variables confirmed that there was a highly significant main effect of level

(F=55.23, p<.001) and complementizer (F=15.22, p<.001), as well as a significant interaction effect (F=13.24, p<.01).

Turning to their judgments by condition, the L2 group judged *wh*-island sentences in the lexical condition (M=3.4, SD=0.8) as slightly more acceptable than those in the pronoun conditions (M=3.3, SD=0.8) although the difference was only marginal (t=1.84, p=.07). Analysis by level revealed that Level 1's rating in the lexical condition was significantly higher than in the pronoun condition (t=2.53, p<.05), whereas the difference was not significant in Level 2 (t=0.17, p=.86) as shown in Table 4.

Table 4. L2 Ratings for *Wh*-Island Constructions by Condition

	Lexical		Pronoun			
	M	(*SD*)	*M*	(*SD*)	*t*	*p*
Level 1 (n=33)	3.3	(0.7)	3.0	(0.6)	2.53	.02
Level 2 (n=32)	3.6	(0.9)	3.6	(0.8)	0.17	.86

A mixed ANOVA with level and condition as variables revealed only the main effect of level (F=6.60, p<.05). The effect of condition was only marginally significant (F=3.37, p=.07), and there was no significant interaction effect between condition and level (F=2.49, p=.12).

The L2 judgment patterns were somewhat different from that of the NSs as illustrated in Figure 2: while the NS group showed a moderate preference for the pronoun condition, the L2 learners, especially Level 1, exhibited a moderate preference for the lexical condition.
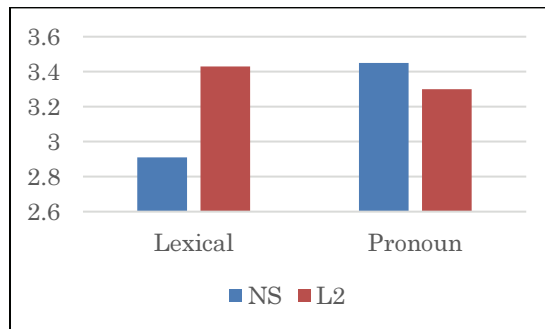

Figure 2. Acceptability ratings for *wh*-island sentences: NS vs. L2

The findings so far can be summarized as follows. First, NSs distinguished long-distance *wh*-movement across *that* from that across *whether*, judging non-island sentences as more acceptable than *wh*-island sentences. L2 learners varied in their judgment patterns depending on their levels. Those who accepted non-violating sentences distinguished *wh*-island sentences from non-island sentences, giving significantly lower ratings to *wh*-island sentences, just like the NSs. In contrast, those who did not accept non-island sentences gave

generally low ratings to both island and non-island sentences.

Second, Level 2 gave higher ratings to the *wh*-construction than Level 1, although Level 2's rating for *wh*-island sentences were lower than their rating for non-island sentences. The finding suggest that their increased proficiency makes *wh*-island sentences—as well as non-island control sentences—more acceptable.

Third, the NS group showed a weak preference for the *wh*-island sentences in the pronoun condition over those in the lexical condition, suggesting sentences involving lower processing costs are more acceptable for NSs. In contrast, L2 learners' ratings for the two conditions were not much different, although the lexical condition received a slightly higher rating in Level 1. The findings suggest that the lower processing costs manipulated by the intervening material did not increase L2 learners' acceptance of *wh*-island sentences.

4.2 RTs

The RT analyses revealed that lexical nouns incurred significantly higher processing costs than pronouns in terms of RT. In this section, we will discuss NSs' and L2 learners' RT patterns, focusing on *wh*-island constructions.

4.2.1 NSs
The NSs' RTs by region in the lexical and pronoun conditions are summarized in Table 5, along with the *t*-test results for each region. Up to R2, RTs in the *wh*-island construction were not significantly different between the lexical and pronoun conditions, which is quite natural in that both conditions had the same lexical material up to R2. At R3, where the matrix subject NP was encountered, RTs were significantly longer in the lexical condition than in the pronoun condition ($t=3.01$, $p<.05$). The same slowdown in the lexical condition was observed at R6, where the embedded subject NP occurred ($t=3.20$, $p<.05$).

Table 5. NSs' RTs by Condition: *Wh*-Island Sentences

Region	Lexical		Pronoun		*t*	*p*
	M	*(SD)*	*M*	*(SD)*		
1	690.5	(330.4)	834.6	(505.9)	-0.83	.43
2	493.3	(147.6)	442.1	(82.9)	1.50	.18
3	574.6	(191.3)	449.4	(172.2)	3.01	.02
4	669.3	(279.6)	566.3	(208.9)	1.92	.09
5	603.1	(198.7)	881.1	(713.3)	-1.20	.27
6	676.5	(220.2)	529.7	(129.0)	3.20	.01
7	524.7	(166.0)	524.1	(195.1)	0.01	.99
8	538.0	(189.4)	602.0	(272.1)	-1.72	.12
9	488.5	(170.7)	540.9	(201.8)	-2.94	.02
10	1776.3	(1299.7)	1705.3	(1024.9)	0.45	.67

The longer RT at the lexical condition in R3 and R6 replicates previous findings that lexical nouns are more costly to process than pronouns (Gibson, 1998; Gordon et al., 2001; Grodner & Gibson, 2005; Warren & Gibson, 2002). The RT difference between the lexical and pronoun conditions at R3 and R6 can be attributed to at least two reasons, as discussed earlier. First, the intervening lexical nouns in the test items such as 'the professor' and 'the assistant' consisted of multiple syllables whereas the pronouns such as 'she' and 'they' had only one syllable. Therefore, the sheer amount of the material was longer in the lexical condition. Also, in terms of their properties, processing new discourse referents might have incurred higher cost than processing pronouns, as proposed by Gibson (1998). These factors might have conspired to increase the RTs at lexical NPs.

Looking at our critical regions, R8 and R9, the NSs were found to slow down in the pronoun condition. At R8, the RT in the pronoun condition (M=602.0, SD=272.1) was longer than that in the lexical condition (M=538.0, SD=189.4), but the difference between them did not reach a level of significance (t=1.72, p=.12). A spillover effect was observed at the following region R9 (t=2.94, p<.05), the RT in the pronoun condition (M=540.9, SD=201.8) being significantly longer than that in the lexical condition (M=488.5, SD=170.7).

The findings are not easy to interpret because previous findings on RT tended to suggest that long-distance filler-gap dependency resolution is more costly when there intervenes more lexical material (Grodner & Gibson 2005; O'Grady, 2008; Roberts & Gibson, 2002). As discussed earlier, long-distance filler-gap integration incurs high processing costs. Sometimes, however, excessive processing demands due to heavy interference might cause processing breakdown, not reaching a successful argument-predicate integration. If a *wh*-island sentence is the cases, processing the intervening lexical NPs as well as other material inside the dependency—including the intervening *wh*-phrase *whether*—and crossing a clausal boundary might have exhausted memory resources and the fronted *wh*-argument might have decayed and thus are not reactivated when the parser has encountered its predicate. As a result, no integration might be attempted, and in turn, little integration cost might be charged at R8. By contrast, pronouns do not charge as much processing costs and thus do not as seriously interfere the reactivation of the *wh*-argument by the time the parser encounters the embedded verb at R8. Therefore, the parser has more chance for the long-distance filler-gap dependency resolution in the pronoun condition, and integration costs are charged.

This interpretation is coherent with the findings on the NS groups' acceptability judgment in the two conditions. Recall that they judged *wh*-island sentences in the pronoun condition as more acceptable than the lexical condition. If we relate sentence acceptability with its processibility, NSs tended to prefer *wh*-island sentences which were more processible. In other words, the pronoun condition allowed more memory resources available for successful interpretation in terms of long-distance argument-predicate

association at R8, leading to higher acceptance of *wh*-island sentences in the pronoun condition.

4.2.2 L2 learners

L2 learners' RTs by condition in *wh*-island sentences are presented in Table 6, along with the results of the pairwise comparison at each region.

Table 6. L2 Learners' RTs by Condition: *Wh*-Island Sentences

	Lexical		Pronoun			
Region	*M*	*(SD)*	*M*	*(SD)*	*t*	*p*
1	819.4	(436.6)	719.7	(309.6)	1.50	.14
2	585.3	(160.8)	634.3	(296.1)	-1.19	.24
3	922.4	(406.8)	578.2	(232.1)	6.86	.00
4	818.0	(344.4)	800.4	(495.7)	0.26	.79
5	867.0	(408.0)	910.4	(420.3)	-0.68	.50
6	766.3	(276.1)	562.9	(182.5)	5.19	.00
7	549.6	(150.9)	505.5	(128.9)	1.81	.08
8	480.7	(141.0)	508.9	(169.6)	-1.57	.12
9	472.6	(152.2)	493.6	(147.1)	-1.21	.23
10	1298.5	(1526.5)	1398.1	(1043.1)	-0.60	.55

At R3 and R6, where intervening NPs varied between lexical nouns and pronouns, RTs were significantly longer in the lexical condition than in the pronoun condition, indicating that learners' processing of pronouns was faster than their processing of lexical nouns. The difference between the two conditions were significant not only at R3 (lexical condition: $M=922.4$ vs. pronoun condition: $M=578.2$) but at R6 (lexical condition: $M=766.3$ vs. pronoun condition: $M=562.9$). The difference between the two conditions were highly significant at R3 ($t=6.86$, $p<.01$) as well as at R6 ($t=5.19$, $p<.01$).

Unlike the RT contrasts at R3 and R6, the difference at R8 was not significant between conditions. While the learners took a little longer in the pronoun condition ($M=508.9$) than in lexical noun condition ($M=480.7$) at R8, the difference was not found to be statistically significant ($t=1.57$, $p=.12$). Similarly, they took slightly longer time at R9 in the pronoun condition than in the lexical condition ($M=493.6$) than in the pronoun condition ($M=472.6$), but the difference was not significant ($t=1.21$, $p=.23$).

Analysis of RTs by level yielded similar results. The two level groups' RTs by region in the lexical and pronoun conditions are summarized in Table 7, along with the results of pairwise comparisons. As shown in Table 7, both Level 1 and Level 2 took longer RT in the lexical condition at R3 and R6. To begin with Level 1, their RT in the lexical condition ($M=1011.6$) was significantly longer than that in the pronoun condition ($M=612.1$) at R3 ($t=5.18$, $p<.01$). Also, their RT in the lexical condition ($M=803.4$) was significantly longer than that in the pronoun condition ($M=565.8$) at R6

(t=3.74, p<.01). On the other hand, their RT difference between the two conditions was not significant at R8 (t=1.62, p=.12).

Table 7. L2 Learners' RTs by Condition: *Wh*-Island Sentences

Region	Level 1			Level 2		
	Lexical	Pronoun	t	Lexical	Pronoun	t
1	895.3	1037.0	-2.38	763.9	908.8	-1.23
2	629.7	617.7	0.33	537.5	652.3	-1.53
3	1011.6	612.1	5.18**	826.3	541.6	4.56**
4	836.1	969.9	-1.29	797.9	612.1	2.63*
5	878.3	1009.1	-1.78	854.5	799.9	0.52
6	803.4	565.8	3.74**	721.9	559.4	4.01**
7	537.4	523.6	0.47	565.6	481.9	2.07*
8	474.5	509.9	-1.62	488.4	507.7	-0.64
9	450.5	469.1	-0.75	496.3	520.1	-0.94
10	1318.1	1279.7	0.22	1275.8	1535.5	-0.88

Note. * indicates p<.05; ** indicates p<.01

A comparable trend was observed in Level 2. They slowed down in the lexical condition at R3 and R6, and spillover effects were found in the following regions R4 and R7. At R3, there was a significant RT difference (t=4.56, p<.01) between the lexical condition (M=826.3, SD=374.4) and the pronoun condition (M=541.6, SD=180.6). A spillover effect followed in R4, showing a significant RT difference (t=2.63, p<.05) between the lexical condition (M=797.9, SD=354.9) and the pronoun condition (M=612.1, SD=199.8). At R6, there was also a significant RT difference (t=4.01, p<.01) between the lexical condition (M=721.9, SD=229.7) and the pronoun condition (M=559.4, SD=164.3). A spillover effect followed in R7, showing a significant RT difference (t=2.07, p=.05) between the lexical condition (M=565.6, SD=178.3) and the pronoun condition (M=481.9, SD=119.8).

Despite the difference in RTs for processing pronouns and lexical NPs, little difference was found at R8 between the lexical condition (M=488.4, SD=157.4) and the pronoun condition (M=507.7, SD=188.2). The lack of difference between conditions at R8 (t=0.64, p=.53) or R9 (t=0.94, p=.40) suggests that the increase or decrease in processing costs in at R3 or R6 did not influence the RT at the critical regions around R8.

Figure 3 demonstrates the two learner groups' time courses while processing *wh*-island sentences in the two conditions.
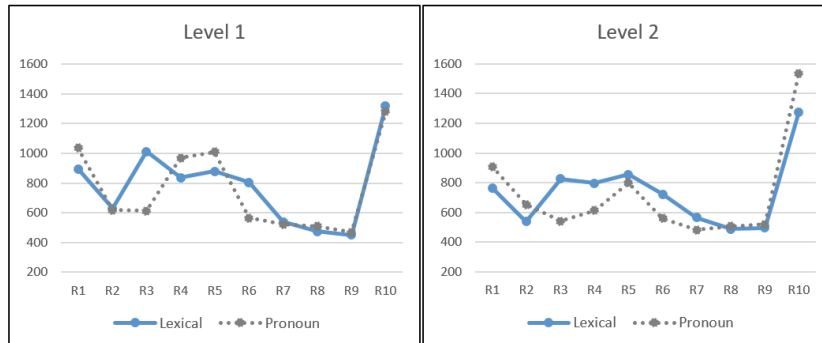
Figure 3. L2 learners' time course in self-paced reading: *Wh*-island sentences

   Learners at the two levels have a few things in common. First, both
levels slowed down at the lexical NPs in R3 and R6, which is comparable to
the one observed in the NS group. Second, no clear difference at the R8—or at
R9—was found between the two conditions. The learners' performances at R8
and R9 are somewhat different from the NSs', who slowed down at R9 in the
pronoun condition.

   On the other hand, the two levels showed a few differences. First, Level
2 required shorter RTs up to R7. Their mean RT per region (M=687ms) was
shorter than Level 1's (M=780ms), although it was longer than the NSs'
(M=604ms). Level 2's shorter RTs at individual regions indicate that their
access to L2 words have been more automatized. From R8, Level 2 tended to
slow down, which suggests that the integration costs were incurred at R8. The
longer RT near the end of the sentence also suggest Level 2 learners' additional
processing for sentence wrap-up, which probably involve completion of
syntactic and semantic interpretation of a sentence. Achieving interpretation,
then, might be related to their higher acceptance of *wh*-island sentences.

4.3 Acceptability judgment and processing costs

Acceptability judgment and RT responses put together, NSs' acceptability of
wh-island sentences correlated with their RTs around the critical region. Their
acceptability ratings were higher in the pronoun condition, and their RT
significantly slowed down in the same condition. The parallelism between
acceptability judgment and processing costs suggests that processing demands
are indeed responsible for the acceptability of highly complex sentences such
as *wh*-island construction. In contrast, L2 acceptability judgment did not show
a parallel relation with their acceptability judgment. While the learners showed
a moderate preference for the lexical condition, their RT neither increased nor
decreased in the same condition. The finding suggests that L2 acceptability
was not affected by processing difficulties in terms of RTs.

   The findings from NS and L2 acceptability judgments are comparable
to those in Hahn (2017). As in Hahn (2017), NSs judged island sentences in

the pronoun condition as more acceptable whereas L2 judged those in the lexical condition as more acceptable. The preference found in our study, however, was much weaker compared with Hahn (2017), which showed a significant rating differences between conditions. The lack of significance in the current study is most probably due to different measuring instruments used in the two studies. In Hahn (2017), judgements were made while the whole sentence string remained on the computer screen. In the current study, judgments were made after self-paced reading, where the lexical content of the previous region disappeared as the parser moves to the next region. In this non-cumulative mode, readers cannot go back to previous regions, and the decisions rely entirely on memory. Due to the non-cumulative nature of the current experiment, the lexical material processed earlier might have deteriorated while parsing, causing the participants to be less decisive in their acceptability judgments.

4.4 The effect of proficiency

Another major finding of the study is that acceptability judgment is affected by learners' proficiency. More proficient learners correctly accepted the grammatical non-island construction and gave a significantly lower rating to the *wh*-island construction, showing that they can distinguish between island and non-island sentences. Still, *wh*-island sentences received higher ratings by more proficient learners than by less proficient learners. The findings suggest that more proficient learners, probably with more efficient sentence processing, are more likely to process the highly complex semantic/structural configuration such as *wh*-dependency resolution. Recall that Level 2 learners' mean RTs up to the critical region was shorter than Level 1's, revealing Level 2's increased processing speed and efficiency.

5 Conclusion

The major findings of this study can be summarized as follows. In acceptability judgment, Level 2 learners, who showed acceptance for non-island sentences, gave significantly lower ratings to *wh*-island sentences compared to non-island sentences, exhibiting a judging pattern comparable to the NSs'. By contrast, Level 1 learners gave ratings near 3.0 to both *wh*-island and non-island sentences. Second, Level 2 learners tended to give higher ratings to *wh*-island sentences compared to Level 1 learners. Third, while the NSs showed a weak preference for *wh*-island sentences in the pronoun condition, L2 learners as a whole showed a weak preference for the lexical condition.

In self-paced reading, all participants required shorter RTs at pronouns compared to lexical NPs at R3 and R6. Second, while the NSs slowed down in the pronoun condition at R9, no such slowdown in any one condition was observed either in Level 1 or in Level 2. Third, Level 2's parsing speed was

faster than Level 1's, confirming that proficient L2 learners' automaticization process is in progress, moving toward more automatic and cost-saving mode.

Despite their gradual change toward automatic processing, Level 2 learners' acceptability ratings in the lexical condition and the pronoun condition displayed a different pattern from those of NSs. Their RTs at the critical region were not affected by different processing costs manipulated by the two conditions.

Taken together, our findings replicates previous findings on NSs' acceptability judgment and processing: acceptance decreases as processing costs rise. Our findings also lend support to the prediction that learners with higher level of proficiency will be more acceptant to *wh*-island sentences. On the other hand, no evidence was found as to the role of processing demands in terms of intervening material on L2 learners' acceptability of *wh*-island. The current paper proposes that L2 learners, including high-level learners, tend to employ a lexical/semantic parsing strategy, unlike NSs, who are known to engage in structure-based parsing. As they rely on lexico-semantic information in resolving the predicate-argument dependency, the lexical information provided by the two intervening NPs could have worked as helping clues rather than interfering obstacles. The lexical/semantic parsing perspectives are in line with Clahsen and Felser (2006), Felser et al. (2012), and Hahn (2017), and somewhat diverge from those of Omaki and Schulz (2011) and Johnson et al. (2016), who claimed that L1 and L2 processing are both structure-based.

The present paper is differentiated from previous works in the Korean L2 literature in that it investigated the relationship between L2 processing and L2 acceptability by directly comparing learners' performance on sentence processing and their performance on acceptability judgment. It employed self-paced reading to measure processing costs and more memory-sensitive acceptability judgment task in the condition where previously parsed words are not available. The idea that L2 learners' higher rating in the lexical condition is related to their lexical parsing strategy still requires further verification through independent future works. The role of processing factors on L2 island effects needs to be investigated through further studies on other types of island constructions. Employing various psycholinguistic tools other than self-paced reading would be desired to verify the findings of this study.

References

Aldosari, S., Covey, L., & Gabriele, A. (2022). Examining the source of island effects in native speakers and second language learners of English. *Second Language Research, 40*(1), 51–77.

Belletti, A., & Rizzi, L. (2013). Intervention in grammar and processing. In I. Caponigro & C. Cecchetto (Eds.), *From grammar to meaning: The spontaneous logicality of language* (pp. 294–311). Cambridge University Press.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In J. Schachter & S. Gass (Eds.), *Linguistic perspective on second language acquisition* (pp. 41–68). Cambridge University Press.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3–42.

Felser, C., Cunnings, I., Betterham, C., & Clahsen, H. (2012). The timing of island in innovative sentence processing. *Studies in Second Language Acquisition*, *34*(1), 67–98.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes, 4*, 93–126.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 94–126). The MIT Press.

Ginzburg, J., & Sag. I. A. (2000). *Interrogative investigations: The form, meaning, and use of English interrogatives*. CSLI Publications.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 1411–1423.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science, 29*(2), 261–290.

Hahn, H. (2017). The effect of processibility on EFL Learners' acceptability judgments of wh-island constructions. *Korean Journal of Applied Linguistics, 33*(4), 51–78.

Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language, 86*, 366–415.

Hofmeister, P., Staum Casasanto, L., & Sag, I. A. (2013). Islands in the grammar? Standards of evidence. In J. Sprouse & N. Hornstein (Eds.), *Experimental syntax and island effects* (pp. 42–50). Cambridge University Press.

Johnson, A., Fiorentino, R., & Gabriele, A. (2016). Syntactic constraints and individual differences in native and non-native processing of wh-movement. *Frontiers in Psychology, 7*. Retrieved from https://doi.org/10.3389/fpsyg.2016.00549.

Johnson, J., & Newport, E. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition, 39*, 215–258.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Kaan, E. (2002). Investigating the effects of distance and number interference in agreement processing: An ERP study. *Journal of Psycholinguistic Research, 31*, 165–193.

Kluender, R. (1998). On the distinction between strong and weak islands: A processing perspective. In P. Culicover & L. McNally (Eds.), *Syntax and semantics, Vol. 29: The limits of syntax* (pp. 241–79). Academic Press.

Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes, 8*, 573–633.

Lakshmanan, U., Kim, K. L., McCreary, R., Park, K., Suen, U., & Lee, S. (2009). L2 learners' sensitivity to strong and weak subjacency-violations in online processing. In M. Bowles, T. Ionin, S. Montrul, & A. Tremblay (Eds.), *Proceedings of the 10th Generative Approaches to Second Language Acquisition C*on*ference* (pp. 136–143). Somerville, MA: Cascadilla Proceedings Project.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*(3), 381–401.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*(2), 1467–1770.

O'Grady, W. (2008). Does emergentism have a chance? In H. Chan, H. Jacob, & E. Kapia (Eds.), *Proceedings of 32nd annual Boston University Conference on Language Development* (pp. 16–35). Somerville, MA: Cascadilla Press.

Ojima, S. (2005). *Theory and evidence in second language research: The acquisition of English by native speakers of Japanese* [Unpublished master's thesis]. University of Essex.

Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition, 33*(4). 563–588.

Pearlmutter, N. J. (2000). Linear versus hierarchical agreement feature processing in comprehension. *Journal of Psycholinguistic Research, 29*, 89–98.

Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In E. Reuland & A. ter Meulen (Eds.), *The representation of (in)definite* (pp. 98–129). MIT Press.

Rizzi, L. (1990). *Relativized minimality*. MIT Press.

Rizzi, L. (2001). Reconstruction, weak island sensitivity, and agreement. In C. Cecchetto, G. Chierchia, & M. T. Guasti (Eds.), *Semantic interfaces: Reference, anaphora and aspect* (pp. 145–176). CSLI Publications.

Roberts, R., & Gibson, E. (2002). Individual differences in sentence memory. *Journal of Psycholinguistic Research, 31*(6), 573–598.

Ross, H. (1987). Islands and syntactic prototypes. In A. Bosch, B. Need, & E. Schiller (Eds.), *Papers from the 23rd regional meeting of the Chicago Linguistic Society* (pp. 309–320). Chicago Linguistic Society.

Shatchter, J. (1989). Maturation and the issue of universal grammar in second language acquisition. In W. Ritchies & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 159–193). Academic Press.

Szabolsci, A., & Lohndal, T. (2017). Strong vs. weak islands. In M. Everaert & H. C. van Riemsdijk (Eds.), *The Wiley Blackwell companion to syntax* (pp. 479–531). John Wiley & Sons, Ltd. Retrieved from https://doi.org/10.1002/9781118358733.wbsyncom 008.

Szabolsci, A., & Zwart, F. (1993). Weak Islands and an algebraic semantics of scope taking. *Natural Language Semantics, 1,* 235–284.

Tabor, W., Villata, S., & Sprouse, J. (2020). Strong and weak islands: A theory of graded accessibility of linguistic domains. *Online proceedings of AMLap 2020.* Retrieved from https://amlap2020.github.io/a/323.pdf.

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience, 2*, 717–726.

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Lawrence Erlbaum Associates, Inc.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition, 85*, 79–112.

White, L. (1992). Subjacency violations and empty categories in second language acquisition. In H. Goodluck & M. Rochemont (Eds.), *Island constraints: Theory, acquisition, and processing* (pp. 445–464). Kluwer.

Hye-ryeong Hahn, Professor
Department of English Education,
Seowon University
377-3 Musimseo-ro, Seowon-gu, Cheongju-si,
Chungcheongbuk-do, 28674, Korea
E-mail: hyerhahn@seowon.ac.kr