

Is Starting Early Beneficial for the Acquisition of English Articles in an EFL Setting?

Sanghee Kim and Mi-Jeong Song*

Kim, Sanghee, & Song, Mi-Jeong. (2020). Is starting early beneficial for the acquisition of English articles in an EFL setting? *English Teaching*, 75(4), 3-32.

This study investigates the role of age of first exposure in the acquisition and processing of English articles in an EFL setting. Fifty advanced Korean learners of English participated in a grammatical acceptability judgment task and a self-paced reading task. The participants were divided into two groups depending on the initial age at which they were constantly exposed to English at least 3 hours per week ('early group' < 12; 'late group' ≥ 12). No significant performance differences were observed between the two groups in the judgment task. However, meaningful differences were found between the reading behaviors of the two groups in the self-paced reading task. These findings show that the learners' age of first exposure has a significant impact on learners' performance when the ability to process English articles in real-time is evaluated. In contrast, it has much less impact on learners' performance on a task which tests explicit knowledge under no time-limit condition.

Key words: English article acquisition, age effect, age of first exposure, EFL learning setting, L2 sentence processing, real-time language comprehension, task effects

*First Author: Sanghee Kim, Graduate Student, Department of Linguistics, The University of Chicago
Corresponding Author: Mi-Jeong Song, Professor, Department of English Language and Literature, Seoul National University; Gwanak-ro 1, Gwanak-gu, Seoul 08826, Korea; Email: mjs@snu.ac.kr

Received 19 August 2020; Reviewed 29 October 2020; Accepted 16 December 2020

1. INTRODUCTION

English articles are known as one of the most “notorious” grammatical features that pose difficulty to L2 learners of English (Jarvis, 2002; Liu & Gleason, 2002; Master, 1987; Murphy, 1997; Tarone, 1985; Thomas, 1989). Even learners who are at an advanced level of English proficiency are reported to make errors in using articles consistently (Butler, 2002; Master, 2002; Park, 2005). As such, it has been a challenge for L2 English to learn and use English articles successfully.

A number of researchers have attempted to investigate the underlying reasons for the difficulty. Five major causes have been proposed: (1) semantic complexity of articles such as concepts as genericity (Ionin, Montrul, Kim, & Philippov, 2011; Ionin, Montrul, & Santos, 2011), specificity and definiteness (Ionin, Ko, & Wexler, 2004), and presuppositionality (Ko, Ionin, & Wexler, 2010); (2) misconception or misjudgment on noun countability (Celce-Murcia & Larsen-Freeman, 1999; Park & Song, 2008; Snape, 2008); (3) influence from learners’ first language (Ionin & Montrul, 2010; Ionin, Zubizarreta, & Maldonado, 2008; Jarvis, 2002; Liu & Gleason, 2002; Robertson, 2000); (4) lack of perceptual saliency of English articles (Goldschneider & DeKeyser, 2001); (5) age of first exposure (DeKeyser, 2000; Song, 2014).

Among the five causes mentioned above, the current study particularly focuses on the role of the age of first exposure. Even though age effect in the acquisition of morphosyntax and phonetics has been extensively examined in the second language acquisition literature (Bialystok, 1997; Birdsong & Molis, 2001; DeKeyser, Alfi-Shabtay, & Ravid, 2010; Flege, Yeni-Komshian, & Liu, 1999; Johnson & Newport, 1989), only few studies explored the acquisition of English articles. Although several researchers suggested that successful acquisition of English articles benefits from an earlier starting age (DeKeyser, 2000; Song, 2014), more evidence is required to support the argument. In addition, little work has been conducted on the relationship between the age of first exposure and the ability to process English articles, and few studies on this issue have been done in EFL settings where extensive, high quality of input is often limited. Therefore, by selecting 50 highly advanced Korean EFL college students who have been exposed to English extensively for more than 10 years and differ in their age of first exposure, this study explores whether their initial age when a minimum of 3 hours of exposure began to be consistently provided influences their acquisition and processing of English articles, utilizing two different tasks.

Tasks that measure learners’ L2 proficiency can be categorized into two types, depending on whether they focus on learners’ product or process. The examples of the first type are untimed grammatical judgment tests, acceptability judgment tests, elicitation tests, cloze tests, spoken narratives and writing tests. The data collected from such tasks can reveal much of learners’ L2 knowledge and their developmental phases of grammatical

forms (Parrish, 1987; Thomas, 1989). The examples of the second type are self-paced reading tasks and eye-tracking tasks. The data collected from such tasks can tap into learners' real-time processing and their instant comprehension of a form (Kim & Lakshmanan, 2008; Trenkic, Mirkovic, & Altmann, 2014). Since both product and process data are crucial in understanding language acquisition (Jiang, 2012; Suzuki & DeKeyser, 2015), the present study investigates the relationship between age of initial exposure and acquisition of English articles with the following two tasks: an untimed grammatical acceptability judgment task and a self-paced reading task. By utilizing two tasks measuring knowledge and use, this study attempts to examine how the effect of age of first exposure will be manifested in each task.

2. BACKGROUND

2.1. English Article System and Second Language Learners' English Article Acquisition

A well-known description on English articles is featured in Huebner's (1983) semantic wheel, which is rooted on the idea of Bickerton (1981). The framework follows that the English article system can be organized by two universal features of referentiality. The two binary features are [\pm Specific Referent] and [\pm Hearer Knowledge].¹ The combination of these features generates four semantic segments to classify English articles: Type 1 [-SR, +HK]; Type 2 [+SR, +HK]; Type 3 [+SR, -HK]; Type 4 [-SR, -HK]. Type 1 is generics, which is marked with *a(n)*, *the*, or no article. For instance, *a model* in *Hannah thought that a model has to be slim* does not indicate a specific referent but an entity that is known to the hearer. Type 2 is referential definites, where a noun is preceded by *the*. For example, *the car*, in *A car suddenly stopped on the road. Sean noticed that the car had a flat tire*, has a specific referent and is an entity known by the hearer. Type 3 is referential indefinites, which is marked by *a(n)*, or no article. Without any information provided, *a girl* in "Yesterday there was a girl screaming and shouting crazily" has a specific entity, but the entity is unknown to the hearer. A first-mentioned noun phrase is falls into Type 3. Type 4 includes nonreferentials such as *a(n)*, or no article. For instance, *a house* in *Jacob wants to buy a house someday* is Type 4 as it does not have a specific referent nor is it known to the hearer. In addition to these four types of articles, a new type was also added to take account

¹ In Huebner's (1983) original work, the two binary features were named as [\pm Specific Referent] and [\pm Assumed Known to the Hearer]. Adopting his work, Thomas (1989) referred to these two features as [\pm Specific Referent] and [\pm Hearer Knowledge]. The feature values are now called by the latter name.

of prefabricated patterns and expressions used as chunks. It was referred to as “conventional use” (Butler, 2002; Park, 2005), or Type 5. Geographical names and building names such as *The Mississippi River* and *The Plaza Hotel* were included for this article type. In consistent with the previous work, the framework of the present study is illustrated by two binary features [\pm SR, \pm HK] and is featured by five types: Type 1, Type 2, Type 3, Type 4, and Type 5.

In terms of English article acquisition by second language learners of English, many studies have reported *the*-flooding by learners (Huebner, 1983; Master, 1987; Parrish, 1987; Thomas, 1989). This *the*-flooding refers to overuse of *the* by the learners in a context where either *a(n)* or *no article* is appropriate. Considering that *the* appears in a [\pm SR, +HK] context, some researchers explained that *the* flooding phenomenon is due to the learners’ tendency to connect [+HK] to *the* and use *the* whenever the learners assume the noun phrase being referred to is known to the hearer. Along with an overuse of *the* by the learners, it has been reported that while *the* is considerably easy to learn, an indefinite article, *a(n)*, is the most challenging for the learners (Master, 1994). Regarding the comparative easiness of the acquisition of *the*, it is suggested that learners capture the [+HK] feature value well beyond other features—this would relate to the cause for *the*-flooding. Unlike [+HK], article use in [-HK] contexts is difficult for learners as can be seen from poor performance for Type 3 [+SR, -HK] and Type 4 [-SR, -HK] (Butler, 2002; Park, 2005; Song, 2014). Subsequent studies reported that Type 5 is also challenging for learners, where article use should be memorized as a formulaic expression and a prefabricated term (Butler, 2002; Park, 2005; Song, 2014).

Previous studies attributed the difficulty in learning English articles to various sources, but further investigation is needed to better understand the underlying cause. Among the possible reasons noted earlier, the present study particularly focuses on the role of advanced EFL learners’ age of first exposure to English in acquiring and processing English articles.

2.2. Age of First Exposure and Learners’ English Article Acquisition

The link between maturational effect and language acquisition derives from an early proposal made by Lenneberg (1967). A hypothesis he formed postulates a time threshold, arguably around puberty, beyond which learning a language cannot be the same as learning it before a critical point of time. Since the hypothesis has been proposed, age effect has been a critical issue in second language acquisition.

Research evidence on the role of age of first exposure in the ultimate level of second language acquisition diverges. From one perspective, the age of first exposure to a new second language is essential in the successful acquisition of the language. Many previous

studies have demonstrated a decline of accurate performance by second language learners (DeKeyser, 2000; DeKeyser, Alfi-Shabtay, & Ravid, 2010; Flege, Yeni-Komshian, & Liu, 1999; Johnson & Newport, 1989), a different way of activation of language representation and processing (Dirix & Duyck, 2017; Meulman, Wieling, Sprenger, Stowe, & Schmid, 2015; Paradis, 2004; Ullman, 2001, 2004) when their age of first exposure falls after a certain time threshold.

However, there also have been studies claiming that age of first exposure is not a determining factor in second language acquisition. Not a few studies have shown that second language learners who learned a new language even after a time threshold can be as fluent and successful as native speakers of the target language (Bialystok, 1997; Bialystok & Miller, 1999). Neurological evidence further demonstrated that language system and neural representation are indistinguishable between native speakers and second language learners even if the learners began learning a new language after a certain time period (Abutalebi & Green, 2007; Green, 2003). As such, views on starting age effect in second language acquisition are still controversial.

Meanwhile, earlier work that focused on second language learners' article acquisition demonstrated an influential role of age of first exposure (DeKeyser, 2000; Song, 2014). These studies showed the benefit that early starters have in the successful acquisition of English articles. In a seminal study by DeKeyser (2000), the relationship between the learners' age of first exposure and the level of success to acquire various morphosyntactic features including articles was tested. DeKeyser (2000) recruited 57 Hungarian learners of English, and they were divided by their age of first exposure; age of first exposure below 16 was grouped as the early group and above 16 was categorized as the late group. In the grammaticality judgment task used, the participants were asked to judge the grammatical accuracy of the presented sentences. The result demonstrated a significant negative correlation between the learners' overall scores and their age of first exposure. An additional finding showed that among the grammatical features he investigated, articles was one of the grammatical features strongly influenced by learners' age of first exposures. However, the finding of the study is not conclusive because only seven individual items of articles were tested in the task.

In order to systematically and thoroughly examine the learners' performance on English articles, Song (2014) utilized a theoretical categorization of the English article system based on Huebner's (1983) semantic wheel. Five types of articles, from Type 1 to Type 5, were tested. Each article type had 20 items, and a cloze task was used. 34 Korean immigrants living in America were recruited for the study. The participants were divided into two groups according to their age of first exposure: the early group (age of first exposure ≤ 12); the late group (age of first exposure > 12). The result showed a negative correlation between the learners' age of first exposure and their performance only for Type

3, Type 4, and Type 5; no statistical correlation was found for Type 1 and Type 2. The sensitivity to the [-HK] feature of Type 3, and the formulaic nature of Type 5 was attributed to the difficulty for the late group. Although her instrument was more rigorous than that of DeKeyser (2000) and her study showed that not all article types were sensitive to learners' age of first exposure, it did not show how learners with different age of first exposure process English articles in real-time, which also constitutes an important issue in understanding the nature of English article acquisition.

2.3. Tasks and Learners' English Article Acquisition

Most previous studies on L2 learners' English article acquisition have utilized tasks which can tap into learners' knowledge and their developmental stages, such as grammatical judgment tests, cloze tests, elicitation tests, and writing tests (e.g., Butler, 2002; DeKeyser, 2000; Park, 2005; Song, 2014). Unlike the studies which focus on learners' product, Kim and Lakshmanan (2008) focused on learners' process of comprehension and investigated how the learners interpret semantic features of English articles in real-time. In their study, both a sentence acceptability rating task and a self-paced reading task were conducted. Korean learners of English at an advanced level, and native speakers of English were recruited for the experiment. While the advanced learners demonstrated a performance similar to the native group in the acceptability task, different processing was detected in the self-paced reading task. The total reading time for the targeted sentences by the learner group was longer than the native control group. The authors explained that the longer reading time is due to the learners' weak processing fluency in constructing semantic and contextual meaning at the instant moment. As such, the observation made on real-time processing of English articles provides further understanding of English acquisition by second language learners (Kim & Lakshmanan, 2008).

3. THE PRESENT STUDY

At the beginning of the review section, it was mentioned that most previous studies that investigated the relationship between the age of first exposure and ultimate level of L2 proficiency have been conducted in naturalistic settings, where sufficient exposure to a target language is guaranteed. Many of the studies have presented findings that show "the earlier is better." But despite the absence of relevant empirical finding, this finding has been overgeneralized in foreign language learning settings (Muñoz, 2010). Therefore, many researchers argue that the age effect in foreign language settings needs to be

explored independently, given the differences in the two learning settings and the significant impact it had on foreign language policies (Muñoz, 2010). Following this rationale, this study investigates the impact of the age of first exposure on the acquisition and processing of English articles by highly advanced Korean learners who have been learning English for almost 10 years in the foreign language learning environment. The results of the present study will provide evidence on whether the initial age of being exposed to English in the foreign language setting influences successful acquisition of English articles.

This study also differs from most of the previous studies in the use of tasks. Except one study conducted by Kim and Lakshmanan (2008) reviewed in the previous section, most studies on L2 learners' English article acquisition used tasks that can measure their knowledge of the form. However, along with a grammatical acceptability judgment task that measures learners' explicit knowledge, the current study also utilized a self-paced reading comprehension task, which allows to measure learners' real-time processing and understanding of English articles. Although Kim and Lakshmanan (2008) used similar tasks, the data analysis of the present study is different. Their study analyzed the total reading time for each sentence. However, the total reading time in fact may not indicate how articles were processed; it may simply demonstrate the learners' reading fluency. Instead, in this study, reading times were measured and calculated for each region of interest. As the focus is the processing of articles but not on reading fluency but on the learner's understanding of English articles, we specifically measured the regions that involve understanding of articles. Moreover, the present study observes how the age of first exposure affects the learners' processing of articles, which was not the focus of Kim and Lakshmanan's (2008) study.

The present study addresses two research questions:

1. How does the learners' age of first exposure affect their performance on the grammatical acceptability judgment task?
2. How do the learners process English articles in real-time in the self-paced reading task, and what is the effect of the age of first exposure in their processing?

4. METHODOLOGY

4.1. Participants

Fifty advanced Korean learners of English were recruited for the experiment. The participants were (under)graduate students at major universities in Seoul. They were

compensated ₩8,000 for their participation. The participants were qualified by two major conditions. First, they should speak Korean as their dominant language. This qualification was to ensure that the participants were Korean learners of English who learned English as their foreign language. Learners who did not learn English in an EFL setting (e.g., who lived abroad during their adolescence) were also excluded from the study. Second, the participants' English proficiency should be at the highest level. Since the purpose of the present study is to examine the effect of age of first exposure on the ultimate level of success (not the rate) in the acquisition and processing of English articles, it was important to select learners who were approaching the final state of L2 acquisition. The participants were asked to submit an official record of an internet-based test of English as a foreign language (iBT TOEFL). Only those with a score of 110 or higher out of 120 were eligible to participate.

The participants were then asked to answer a language background questionnaire before the main experiment. The participants were requested to elaborate their experience to study English and live in an English-speaking country, the amount of English input they received, and the age to start to learn English and the learning methods. Six participants were excluded as they did not provide sufficient information on their language profile. In the main analysis, data from 44 participants were used.

As the participants were not informed of the definition of "age of first exposure" in the questionnaire, it was manually coded based on the learners' responses. The age of first exposure was determined based on Larson-Hall (2008)'s claim that there should be at least 3 hours of a minimum amount of consistent input or exposure for a L2 learner to acquire a language in a foreign language setting. For example, many of the participants answered that they started to learn English in the 3rd grade (10 years old) in the primary school (40 minutes x 2 sessions per week), but their age of first exposure was marked only when they were reported to be involved in other English extracurricular activities that amounts to at least 3 hours per week. In addition, a response such as "I started to learn English at 5. But at that time, I only memorized English alphabets" or "I was taught English when I was in my mom's womb" was not categorized as early exposure group.

Based on the learners' responses and the criterion, the participants were grouped as the "early group" (age of first exposure < 12) and the "late group" (age of first exposure ≥ 12). The criterion of the age division in the current study adopted the standard used in many earlier studies (e.g., McDonald, 2006; White & Genesee, 1996). Participant information of each group is presented in Table 1.

TABLE 1
Language Background Information

	Early Group ($n = 22$) (AFE < 12)	Late Group ($n = 22$) (AFE \geq 12)
Mean age (years)	24.73	24.91
Sex (female)	18	22
Mean age of first exposure (years)	6.36	12.64
Mean iBT TOEFL score (out of 120)	114.59	114.09
Mean length of residence abroad (years)	2.81	1.71

Note. AFE = Age of first exposure.

A Welch's t-test showed that the early group and the late group did not differ in the level of proficiency ($p = .539$) and the length of residence abroad ($p = .215$). The absence of difference verified that the two groups were comparable in language experience and in the current language ability. The only difference was their age of first exposure to English ($p < .001$). These analyses showed that no other confounding factors intervened, and the impact of age factor on the performance of the two age groups was explored.

4.2. Materials

Each item consisted of a context sentence and a target sentence (Table 2). The two sentences were designed to function as a single discourse context ensemble. The first sentence was a context sentence, which provided background information for the use of articles in a target sentence. The second sentence was a target sentence, which contained the targeted noun phrase. The noun phrase had different types of articles, from Type 1 to Type 5, which were selected based on the categorization of English article system used in the present study. Different from the prior work is the inclusion of weak definites into Type 5. Weak definites include noun phrases that are expressed with *the* but do not indicate a specific entity (e.g., Carlson & Sussman, 2005). For instance, *the newspaper* in "read the newspaper" has no specific *newspaper* being referred to, but the bare noun in the noun phrase is preceded by *the*. Given the formulaic characteristic of their use, weak definites were categorized as article Type 5 in this study.

TABLE 2
A Sample Set of the Experimental Material

Condition	Context Sentence	Target Sentence
Type1 [-SR, +HK]	Generics: <i>a(n)</i> , <i>the</i> , \emptyset	
Acceptable	There may be some stereotypes on careers.	Hannah thinks that <u>a model</u> has to be slim.
Unacceptable	There may be some stereotypes on careers.	Hannah thinks that <u>model</u> has to be slim.

Type 2 [+SR, +HK] Referential definites: <i>the</i>		
Acceptable	A woman suddenly fainted on the road.	Doctor Lee reported <u>the state</u> of the woman's health.
Unacceptable	A woman suddenly fainted on the road.	Doctor Lee reported <u>a state</u> of the woman's health.
Type 3 [+SR, -HK] Referential indefinites, first mentions: <i>a(n), \emptyset</i>		
Acceptable	Graduation ceremonies were held in February.	Judy gladly received <u>presents</u> from her close friends.
Unacceptable	Graduation ceremonies were held in February.	Judy gladly received <u>the presents</u> from her close friends.
Type 4 [-SR, -HK] Nonreferentials: <i>a(n), \emptyset</i>		
Acceptable	Some people find joy by helping others.	Olivia has been <u>a nurse</u> most of her life.
Unacceptable	Some people find joy by helping others.	Olivia has been <u>the nurse</u> most of her life.
Type 5 Conventional use: <i>the, \emptyset</i>		
Acceptable	There are famous hotels in New York.	Charlie stayed at <u>the Plaza Hotel</u> with his girlfriend.
Unacceptable	There are famous hotels in New York.	Charlie stayed at <u>Plaza Hotel</u> with his girlfriend.

The grammatical acceptability judgment task and the self-paced reading comprehension task used the same materials that consisted of 69 items in total: 12 items for Type 1; 18 items for Type 2; 9 items for Type 3; 18 items for Type 4; 12 items for Type 5. Each target item was varied by two acceptability conditions, either “acceptable” or “unacceptable.” The variation on the acceptability condition was made by the presence or absence of the article, or by alternation of the article from *a(n)* to *the* or from *the* to *a(n)*. Accordingly, a total of 138 sentences (= 69 items x 2 conditions) were created for the target sentence. The sentence that functioned as a context sentence remained the same regardless of the acceptability condition of the target sentence.

4.2.1. Norming study

A norming study was conducted prior to the main experiment in order to ensure that the acceptability condition was plausibly designed. Twenty-five native speakers of English were recruited (mean age =24.08; 18 females) from Amazon Mechanical Turk. The participants were asked to rate the sentences by the degree of acceptability. They were instructed to consider the two given sentences—a context sentence and a target sentence—as being introduced within a single context, and to decide how grammatically acceptable the target sentence was in a given context. The participants gave scores in a six-point Likert scale ranging from 1 (“unacceptable”) to 6 (“acceptable”). A total of 207 target sentences (= 69 target sentences x 2 acceptability conditions + 69 context sentences) were used.

A Mann-Whitney U Test was run by using the *wilcox.test* function implemented in the stats library in R (R Core Team, 2016) to test whether the mean ratings for the two conditions significantly differ. The analysis showed that the acceptable condition was rated statistically higher than the unacceptable condition for all five article types ($ps < .001$).

4.2.2. Main experiment

The same material used in the norming study was used in the main experiment. In the acceptability judgment task, all the words in each sentence appeared at once on the computer screen. The sentences in the self-paced reading task were presented word-by-word or phrase-by-phrase (Table 3). The number of regions ranged from six to seven for the context sentence, and eight to nine for the target sentence. The critical word position for Type 1 – Type 4 was Region 4 in the target sentence, where a noun phrase with the target article appeared. The critical word position for Type 5 was Region 4 and Region 5. For instance, “the Plaza Hotel” was presented as “the Plaza / Hotel” for the acceptable condition, and “Plaza / Hotel” for the unacceptable condition. The following two regions after the critical region were the spill-over 1 region, and the spill-over 2 region, respectively. The assignment of the spill-over regions was to reflect a widely recognized characteristic in a self-paced reading task that reading time difference may not be exactly detected at the targeted position alone but may lag through the ensuing regions (Bertram, Hyönä, & Laine, 2000).

TABLE 3
An Example of the Experimental Item for the Self-paced Reading Task

	Region 1	Region 2	Region 3	Region 4
Context	Some	people	find	joy
Target (Accpt.)	Olivia	has	been	a nurse
Target (UnAccpt.)	Olivia	has	been	the nurse
	Region 5	Region 6	Region 7	Region 8
Context	by	helping	others.	
Target (Accpt.)	most	of	her	life.
Target (UnAccpt.)	most	of	her	life.

Note. The targeted noun phrases are bold-faced. “Accpt.” is the “acceptable” condition; “UnAccpt.” is the “unacceptable” condition.

Unlike the words in other regions, the critical region was presented in a phrase-by-phrase fashion. The change was to reflect a behavioral characteristic that participants read or press buttons—the button designated to move on to the next region—fast for highly frequent but short words such as articles and prepositions, and even faster than their normal reading speed for those words (Jiang, 2012). In this respect, the noun phrase at the critical

region was given in a phrase-by-phrase presentation in order to detect participants' processing of articles embedded within the noun phrase.

The 138 item sets were divided into two lists, such that each list contained only one version of each item. A total of 144 filler sentences of different syntactic constructions but similar in complexity and length were included in each target list. Filler sentences also consisted of a context sentence and a target sentence to be comparable to main experimental sentences in the number and format of the presentation. Each list that included both the main and the filler sentences was pseudo-randomized by using the IbexFarm experimental software (<https://spellout.net/ibexfarm/>).

4.3. Procedure

First, the participants completed a language background questionnaire. Then, a self-paced reading task was administered. For the self-reading task, an Arial in a 24-point font size was used for the words presented on the screen. Each item was displayed in two lines, where a context sentence was in the first line and a target sentence was in the second line. Participants were informed to consider the two sentences together to construct a single context. The sentences were presented word-by-word for all regions but the critical region, which was shown phrase-by-phrase. Each region was veiled behind a series of dash marks. Whenever the participants pressed the spacebar, the dashes revealed their hidden words in a non-cumulative moving-window fashion. Reading times (RTs) for each region were measured.

After the participants read each item in the self-paced reading task, complete sentences in each item appeared on the computer screen along with a question, "How contextually grammatical or acceptable is the second sentence?" A Likert-scale ranging from 1 ("unacceptable") to 6 ("acceptable") was presented on the screen. The participants were asked to press a number on their keyboard that corresponded to the score they thought. They were reinstructed to rate how grammatically acceptable the target sentence given the context sentence. There was no time constraint on responding to the question since the purpose of the task was to test the participants' explicit knowledge on English articles.

5. RESULTS

5.1. Grammatical Acceptability Judgment Task

The mean rating scores in the grammatical acceptability judgment task are summarized in Table 4. A Mann-Whitney U Test was implemented by using *wilcox.test* used in an R

environment (R Core Team, 2016) for a non-parametric test. The rating score was compared between the early group and the late group. The acceptable condition was scored significantly higher than the unacceptable condition for both the early group and the late group ($p < .001$). Moreover, both the early and the late group rated article types differently depending on the acceptability condition except for Type 5 (Type 5, $p = 0.64$) (Table 4).²

TABLE 4

A Summary of the Grammatical Acceptability Judgment Task Result

Type	Group	Acceptable	Unacceptable	<i>p</i>
Type 1	Early	4.60 (0.05) ^a	3.61 (0.05)	<.001
	Late	4.71 (0.04)	4.02 (0.05)	<.001
Type 2	Early	5.08 (0.03)	3.73 (0.05)	<.001
	Late	5.01 (0.03)	3.98 (0.05)	<.001
Type 3	Early	4.80 (0.05)	4.22 (0.06)	<.001
	Late	4.68 (0.05)	4.43 (0.05)	<.001
Type 4	Early	4.58 (0.04)	3.48 (0.04)	<.001
	Late	4.76 (0.04)	3.74 (0.05)	<.001
Type 5	Early	4.64 (0.05)	4.10 (0.05)	<.001
	Late	4.56 (0.05)	4.60 (0.05)	<i>n.s.</i>

Note. Mean (standard error) score. No statistical significance marked as *n.s.*

^a Score range from 1-6.

First, the analyzed result indicates that both the early and the late group made correct judgment on the acceptability condition. Second, the insignificant difference for Type 5 by the late group demonstrates that the learners in the late group were insensitive to the difference between the acceptable and unacceptable conditions. This result was different from the early group participants, who were able to tell the difference between the two acceptability conditions for Type 5.

5.2. Self-Paced Reading Task

Reading time data from each region were collected. Instead of the measured data, the residual reading time (RRT) was used for the main analysis. The calculated RRT fixes any reading time differences caused by the length variation of a sentence or a word; RRT is the reading time corrected by word length (Ferreira & Clifton, 1986). Using the RRT, data points that exceeded 3 standard deviation from the overall mean were removed (Jiang, 2012). This procedure was to eliminate influential data points.

Considering that articles are the focus of observation, analysis on the reading time data

² We also compared the score between the two groups within the same type of the same grammatical condition. For the acceptable condition, there was no significant difference between the early and the late group for all article types except for Type 4 ($p < .05$). For the unacceptable condition, there was a significant difference between the two exposure groups for all article types ($ps < 0.05$).

only from the critical region may not fully reflect how articles are processed. Thus, the reading time data in the following spill-over regions were thus included and collapsed into a single data along with the data from the critical region (Table 5). The three regions were referred to as the regions of interest (ROIs), and data points from the ROIs were used for the initial analysis.

TABLE 5
Combined Mean Residual Reading Times (Standard Error) at the ROIs

Type	Group	Acceptable	Unacceptable
Type 1	Early	-8.67 (18.80)	35.40 (18.96)
	Late	38.37 (18.24)	-18.89 (20.61)
Type 2	Early	-34.39 (13.64)	5.10 (16.33)
	Late	23.26 (13.54)	-15.68 (16.88)
Type 3	Early	13.33 (22.19)	21.80 (24.14)
	Late	-16.76 (24.05)	-7.34 (27.24)
Type 4	Early	-32.98 (15.48)	28.02 (14.57)
	Late	-2.72 (14.87)	-17.87 (18.51)
Type 5	Early	-33.91 (17.17)	26.69 (17.16)
	Late	47.12 (16.30)	-29.46 (19.92)

Note. The numbers are in milliseconds.

For data analysis, a Linear Mixed Effects Regression (LMER) analysis (Baayen, Davidson, & Bates, 2008) was built by using an lme4 R package (Bates, Mächler, Bolker, & Walker, 2014).³ In the initial model, the effects that were aimed to be observed were selected as fixed effects: the acceptability condition, the age of first exposure condition, and the interaction of the two conditions. Subjects and items were selected as random effects. The models produced coefficients, standard errors, and *t*-values for the selected effects. The coefficients are the slope of the model, which demonstrate how predictable the measured values are. The standard errors show the certainty of coefficient estimates. The *t*-values present the statistical significance of the coefficients. The coefficients were considered statistically significant if the absolute value of *t* exceeded 2 (Baayen, 2008). In addition to the regression analysis, a planned contrast of pair-wise comparison was conducted. The Tukey test was used with a *multcomp* package in R (Hothorn, Bretz, &

³ The reason to choose this type of regression model was to consider both by-subject analysis from the data averaged over participants (F_1) and by-item analysis from the data averaged over items (F_2) at the same time. While analysis of variance (ANOVA) also captures the significance of main effects by $F_1 \times F_2$ criterion, the interpretation of the analysis becomes difficult in a case where either type of analysis does not reach a statistically significant level. It also fails to handle random variance of subjects and items at the same time. LMER analysis, on the contrary, makes it possible to capture the two types of variance within a single model. For readers who are more familiar with ANOVA, the random variance in ANOVA is comparable to the random effects in the LMER models and main effects in ANOVA to fixed effects in the regression models (Cunnings, 2012).

Westfall, 2008). The analysis on the ROIs was done separately for each type of article. Statistical analyses are given in Table 6.

TABLE 6
Linear Mixed Effect Regression Results at the Regions of Interest

	Coefficient	<i>se</i>	<i>t</i>
Type 1			
(Intercept)	-8.67	19.35	-0.45
Age of First Exposure (AFE)	47.07	27.43	1.72
Acceptability	47.07	27.04	1.63
AFE:Acceptability	-101.38	38.29	-2.65*
Type 2			
(Intercept)	-33.92	17.19	-1.97
Age of First Exposure (AFE)	56.73	24.48	2.32*
Acceptability	38.03	21.48	1.77
AFE:Acceptability	-74.85	30.29	-2.47*
Type 3			
(Intercept)	13.87	26.04	0.53
Age of First Exposure (AFE)	-31.13	35.95	-0.87
Acceptability	7.40	34.89	0.21
AFE:Acceptability	2.93	49.25	0.06
Type 4			
(Intercept)	-32.98	16.16	-2.04
Age of First Exposure (AFE)	30.26	22.47	1.35
Acceptability	61.00	22.47	2.72*
AFE:Acceptability	-76.15	31.75	-2.40*
Type 5			
(Intercept)	-33.91	18.05	-1.88
Age of First Exposure (AFE)	81.01	25.40	3.19*
Acceptability	60.56	25.00	2.42*
AFE:Acceptability	-137.06	35.33	-3.88*

Note. The coefficients were considered statistically significant if the absolute value of *t* exceeded 2.

For Type 1, no main effects were found for the AFE condition and the Acceptability condition; however, a statistically meaningful interaction between the two conditions was found. This indicates that the reading time difference due to the acceptability condition in Type 1 is influenced by the age of first exposure condition. The early group spent longer reading times for the unacceptable items, whereas the late group spent longer times for the acceptable items.

For Type 2, a main effect of the AFE condition was found, but no main effect for the Acceptability condition was found. There was also an interaction between the two main factors. The interaction suggests that the reading time difference caused by the acceptability condition was affected by AFE condition. The early group spent more time reading the unacceptable items than the acceptable items, but the late group spent more time for acceptable items than for unacceptable ones.

For Type 3, the coefficients did not reach a significant level for any of the fixed effects,

demonstrating that the acceptability conditions were not processed differently, and the processing of the articles was not influenced by the AFE condition. No significant interaction was found between the two fixed effects, showing an absence of influence of the AFE condition on the acceptability condition.

For Type 4, a main effect of the acceptability condition was found, but no main effect for the AFE condition was found. There was also a statistically significant interaction between the two main factors. The interaction suggests that the reading time difference caused by the acceptability condition was affected by AFE condition. The late group spent less reading time for the unacceptable condition compared to the early group. A post-hoc Tukey contrast demonstrated that the early group spent longer reading times for the unacceptable condition than for the acceptable condition ($estimate = 61.00, se = 22.47, z = 2.715, p < .05$).

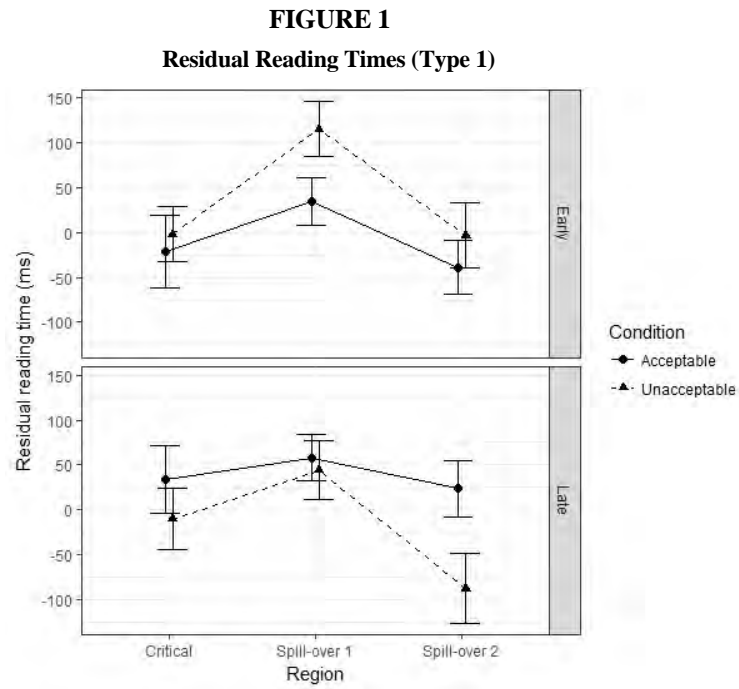
For Type 5, main effects were found for the AFE condition and Acceptability condition, and a statistically meaningful interaction between the two conditions was also found. This indicates that the reading time difference due to the acceptability condition in Type 5 is influenced by the age of first exposure condition. The analysis demonstrates that the late group read the unacceptable condition faster compared to the early group. Meanwhile, interestingly, a post-hoc Tukey pair-wise contrast showed that the late group read the unacceptable faster than the acceptable condition ($estimate = -76.505, se = 24.959, z = -3.065, p = .012$). This is a reverse effect in that more time was needed for the late learners to process the acceptable items compared to the unacceptable items.

In addition to the analyses on the combined reading time at the ROIs, further analyses were conducted on the reading time data at the critical, spill-over 1, and spill-over 2 regions. Similar to the main analysis, an LMER model (Baayen et al., 2008) was used under an lme4 R package (Bates et al., 2014). The regression model was built differently from the initial model. The data of the two AFE groups were separately analyzed, and only the acceptability condition was included as the fixed effect. Item and subject factor were both included as the random effect. The models produced estimates, standard errors, and t -values. Following Baayen (2008), the coefficients were considered to reach a statistically significant level when the absolute t -value exceeded 2.

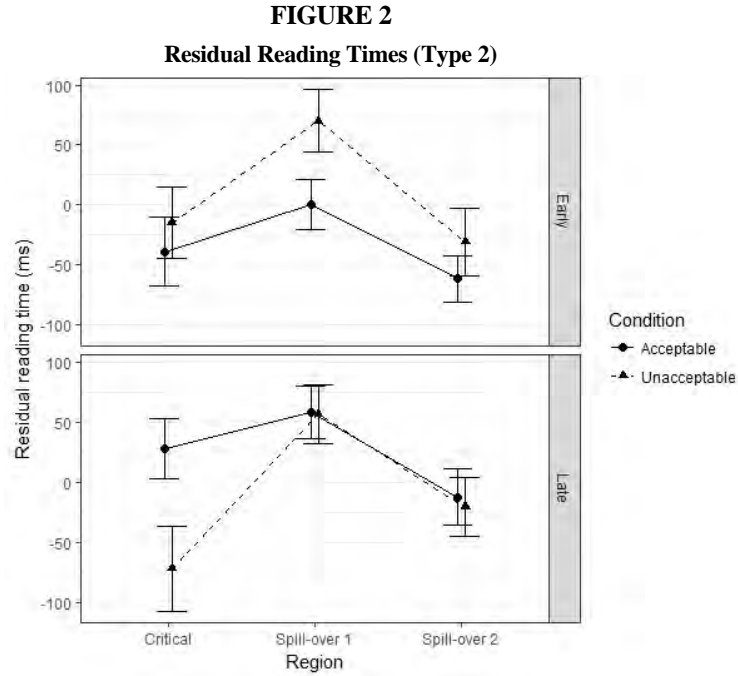
For Type 1, there was main effect of acceptability condition at the spill-over 1 region for the early group ($estimate = 80.87, se = 40.32, t = 2.01$). There was also a main effect of acceptability at the spill-over 2 region for the late group ($estimate = -112.14, se = 52.23, t = -2.03$) (Figure 1).⁴ The unacceptable condition was read slower for the early group while faster for the late group. No statistical differences between conditions were found at the

⁴ We present the residual reading times by exposure group in a separate facet grid in the figures for the sake of the presentation. However, note that both exposure condition and acceptability condition were included as fixed effects in the actual data analysis.

critical region for both groups.



For Type 2 (Figure 2), a main effect was found at the spill-over 1 region for the early group (*estimate* = 71.31, *se* = 35.79, *t* = 2.00), and at the critical region for the late group (*estimate* = -96.56, *se* = 46.95, *t* = -2.06). There was no difference between the two groups in the spill-over 2 region.



RRTs for Type 3 are presented in Figure 3. Despite the trend shown in the figure, regression models for Type 3 showed that no statistical significance was found between the conditions. For Type 4 (Figure 4), a main effect was observed at the critical region for the early group ($estimate = -92.10$, $se = 44.14$, $t = 2.09$). No reading time difference was found by the late group. RRTs for Type 5 are plotted in Figure 5. The regression models showed a statistical difference between the two acceptability conditions by the late group at the critical region late group ($estimate = -131.45$, $se = 45.02$, $t = -2.92$) but not at the other regions. The reading time difference between the two acceptability conditions did not reach a statistical level at any of the regions by the early group.

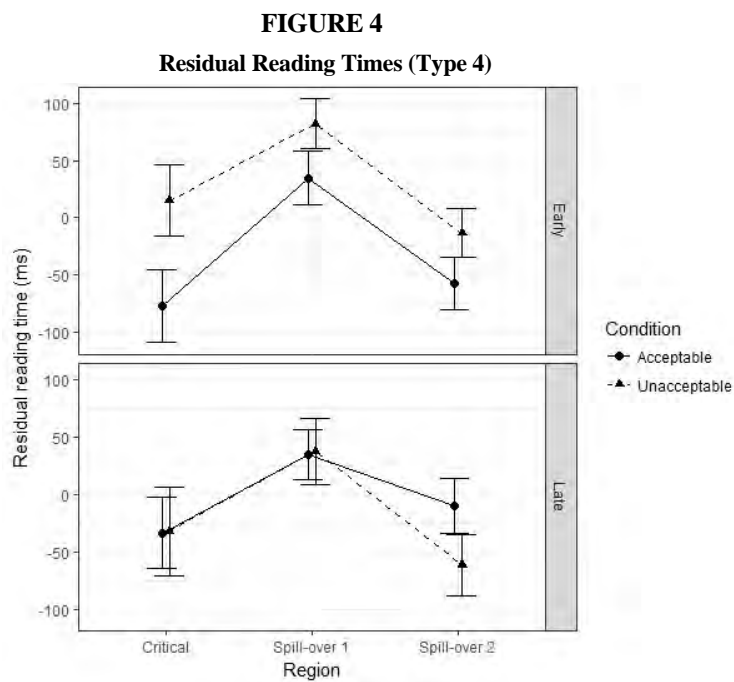
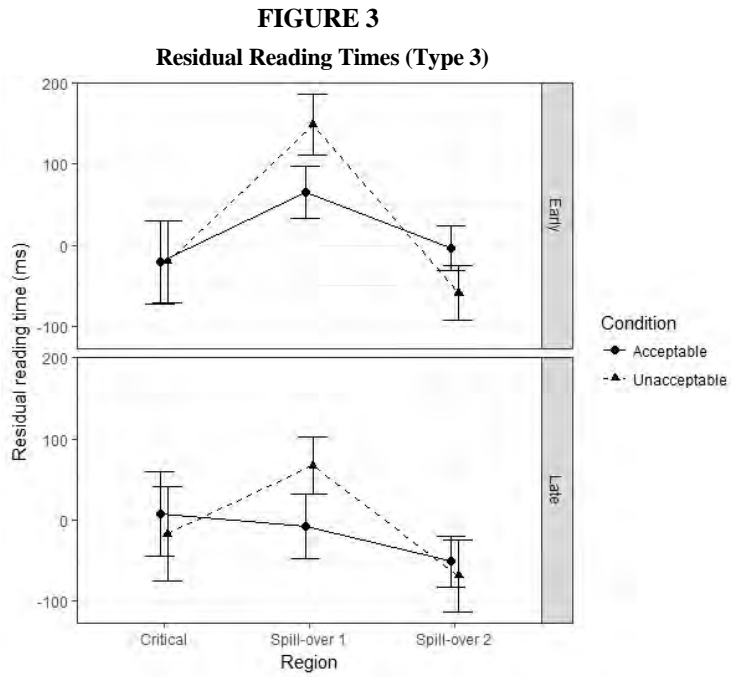
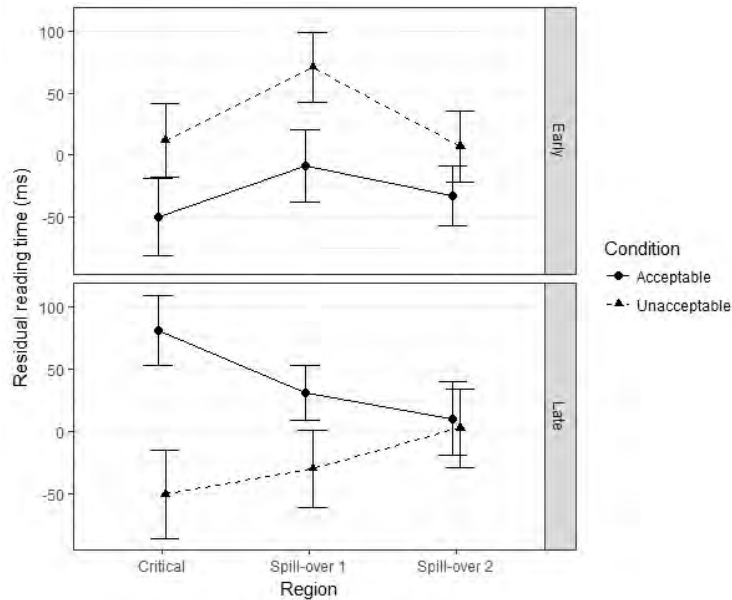


FIGURE 5
Residual Reading Times (Type 5)



Taken together, the results demonstrate that the early and the late group showed a different way of processing English articles in a number of ways.

6. DISCUSSION

The goal of the present study was to examine the role of the age of first exposure in English article acquisition and processing by Korean EFL advanced learners of English. Data from the two tasks, the grammatical acceptability judgment task and the self-paced reading task, were obtained and analyzed.

The performance between the early group and the late group did not differ much in the judgment task. The rating score between the two acceptability conditions showed a statistical difference in both groups for all of the article types except Type 5, for which the late group's performance did not differ. The participants' overall outstanding performance in the article acceptability judgment test is surprising given the previous studies that reported the difficulty of acquiring English articles by second language learners of English (Butler, 2002; Murphy, 1997). The different result can be explained by the strict control of the learners' level of English proficiency in the present study. As noted, the participants' English ability was at highly advanced level. In this vein, a plausible reason for the

Is Starting Early Beneficial for the Acquisition of English Articles in an EFL Setting?

participants' performance in the four types of English articles can be attributed to their high level of English (Kim & Lakshmanan, 2008). This finding also shows that EFL learners' age of first exposure does not affect their acquisition of explicit knowledge in the four types of English articles. That is, the result indicates that even late learners can reach a high level in the acquisition of the English article system in EFL settings.

Yet, as the two groups' performance for Type 5 differed, the age of first exposure (AFE) effect in English article acquisition cannot be completely denied. Although the two age groups performed well for Types 1 [-SR, +HK], 2 [+SR, +HK], 3 [+SR, -HK], and 4 [-SR, -HK], the late group had difficulty making a correct distinction between the acceptable and unacceptable conditions for Type 5, while the early group did not. Such a finding indicates that the early group may benefit from their earlier exposure to English in acquiring Type 5 articles. Type 5 is different from the other article types in that it is used and learned as a chunk or a formulaic expression. As this type is item-based (not rule-based) and is known to be acquired implicitly, the exposure or the familiarity to the words used with Type 5 articles at an early age may contribute to the success of acquiring it. The result is also in line with the previous research finding from the grammatical acceptability task that the learnability of Type 5 is highly influenced by learners' age of first exposure (Song, 2014; Ullman, 2001, 2004). In short, the different performance between the early and the late group for article Type 5 reveals that the initial age of receiving a minimum of three hours of consistent exposure to English affects learners' acquisition of explicit knowledge regarding conventional use of English articles.

While the two age groups performed similarly in the judgment task, they showed a number of different behaviors in processing English articles in the self-paced reading task. As reported in the result section, the analyses on the reading times showed a significant interaction between the acceptability condition and the AFE condition for all article types but Type 3 [+SR, -HK]. This suggests learners' understanding of the correct use of English articles is modulated by AFE. The total reading time data at the ROIs and the RRTs showed that the early exposure group spent longer time reading the unacceptable test items than the acceptable test items, which is an expected behavior. On the contrary, the late exposure group read the unacceptable ones quicker than the acceptable ones.

Meanwhile, what we refer to as the "reverse effect" is observed, where learners spend longer reading times for the acceptable condition rather than the unacceptable condition. The effect is the opposite from the plausibly expected reading performance, where it takes longer reading time for the unacceptable condition and shorter for the acceptable condition. The results showed that in some article types, the late group spent longer time processing the correct version of testing items as opposed to the early group. For example, statistical analyses on RRTs showed that the late group spent longer reading times for the acceptable condition at the spill-over 2 region for Type 1 [-SR, +HK], at the critical region for Type 2

[+SR, +HK], and at the critical region for Type 5. A post-hoc analysis at the ROIs for Type 5 also demonstrated that more reading times were spent by the late group for the acceptable condition than the unacceptable condition. Since a reverse reading time effect was observed throughout the items by the late age group, the interaction effect should be interpreted with caution.

There are two possible interpretations for this reverse effect. One plausible reason comes from the presence of articles. A salient similarity of the items conditioned “acceptable” for Type 1 [-SR, +HK] and Type 5 is that they have articles preceding the noun phrase. For instance, the acceptable condition for the Type 1 [-SR, +HK] noun phrase is *a model*; the unacceptable condition is *model*. This is also the case for Type 5, where the acceptable condition is *the Plaza Hotel*, whereas the unacceptable condition is *Plaza Hotel*. The difference between the two acceptability conditions in both article types is the presence or absence of English articles. In line with the reported reading time result from the late group and the commonality of the acceptable condition, it can be proposed that the reverse effect may be highly contingent on the presence of articles in the acceptable condition. A possible reason for the late group’s spending more time to read items with articles may be related to the characteristic of Korean, the participants’ first language. The absence of an article system in their L1 itself can cause a problem for the late group. As an article is a difficult and an unfamiliar linguistic feature that is absent in the participants’ first language, the presence of an article itself could generate a processing burden and affect article processing. This would explain why the reverse effect was only found among the late group.

The other possible account is concerned with the participants’ incorrect one-to-one mapping of the use of the correct article in the given context. In an EFL setting, learners are explicitly instructed as a rule that *the* is used for repeating a noun phrase that has already been introduced in the immediate context. The explicit instruction also happens for the article, *a(n)*: If a noun phrase appears for the first time, use *a(n)*. Although this rule-like explanation works well in certain contexts, the instructed rule can then be wrongly understood that whenever a noun phrase headed by *a(n)*, it (only) refers to an entity that is contextually new and uninformed to the reader or the listener, and that whenever a noun phrase is headed by *the*, it (only) has an anaphoric function that refers something in the given context. The instructed uses of articles would thus lead to a misleading and erroneous interpretation of the articles.

The erroneous mapping of articles particularly provides a plausible explanation for the late group’s performance for Type 2 [+SR, +HK]. While both acceptability conditions have an article preceding a noun phrase in Type 2, the acceptable condition is headed by *the* while the unacceptable condition is headed by *a(n)*. Relating to the erroneous mapping of articles, it can be interpreted that the late group started to search for a plausible entity that would link the [the NP] as they encountered *the*. In this sense, a longer reading time

for the acceptable condition by the late group may be due to an instructed perception of *the* and *a(n)*.

Another interesting finding comes from the learners' performance on Type 2 [+SR, +HK]. Regardless of the acceptability condition, the late group spent longer time than the early group on average. This result can be explained by a co-referring cost deriving from co-referring process. Concerning the co-referring process in real-time, it has been widely reported that there is a processing cost to link the [the NP] to a previously mentioned noun phrase (e.g., Gordon, Hendrick, & Johnson, 2001). Longer reading times can thus be taken as a processing burden due to a retrieval process, in which [the NP] form is under the processing of being associated with the noun mentioned earlier. Recall that an example item of Type 2 [+SR, +HK] was, "[Context sentence] A car suddenly stopped on the road. [Target sentence] Sean noticed that the car had a flat tire." When it comes to processing a noun phrase headed by *the* as in *the car*, participants should retrieve from their memory what [the NP] refers to. In this respect, the result suggests that the late group is less skilled in linking the noun phrase with *the* to a previously mentioned noun in the given context and may have "shallowly" (Clahsen & Felser, 2006) parsed the referents. This explains why the reading times were longer for the late group.

For Type 4 [-SR, -HK], both the early and the late group demonstrated a sturdy sensitivity to the acceptability condition in the acceptability judgment task. However, in the self-paced reading task, RRT data showed that only the early group but not the late group had the sensitivity to the acceptability condition. The early group read longer times for unacceptable items on the critical region, whereas the late group's reading time between the two acceptable conditions did not differ on all 3 regions. The performance gap can be explained by the learners' different processing of semantic features of Type 4 [-SR, -HK]. It has been shown in earlier studies that the failure to successfully understand and learn English articles comes from the difficult semantic feature of articles. In specific, it was demonstrated that [-HK] feature values is particularly difficult for the Korean learners (Song, 2014). Considering the results of earlier studies with tasks focusing on learners' product data, the processing difference can be explained by the difficulty that the late group has for accurate interpretation of semantic features in real-time.

In addition to the issue on the reverse effect, a point of interest is the absence of main effects and interaction effects for Type 3 [+SR, -HK]. Neither the early group nor the late group was sensitive to the acceptability condition, and no statistical difference were found between the reading times of the acceptability conditions. The reason can be found in the subtle semantic feature of Type 3 [+SR, -HK]. The subtleness is well attested in the norming study where even native speakers of English did not show a significant contrast between the judgments on the two acceptability conditions. The lack of ability to tell from the two conditions by both age groups resonates with previous studies, which demonstrated

that Type 3 [+SR, -HK] is a type of article which highly advanced Korean learners of English have difficulty mastering (Park, 2005; Park & Song, 2008; Song, 2014).

As for the two different tasks, no prominent effect of age of first exposure was found in the grammatical acceptability judgment task; however, a number of reading times data analyses demonstrated that processing differences exist between the two age groups, and the participants' age of first exposure plays an important role in their online processing of English articles.

Two possible interpretations can be suggested for the reasons why different tasks yielded different results. First, the observed gap between the two tasks can be explained by Processing Deficit Approach (Jiang, 2004). This approach explains why the late age group exhibits performance difference in the two tasks. Processing Deficit Approach posits that L2 learners do not fail to have a full representation on the morphology of the target language. According to the Processing Deficit Approach, the late group's performance from the acceptability judgment task demonstrates that the learners were aware of the differences between the two acceptability conditions for most article types. This result indicates that the late learners had a sturdy representation of semantic features and linguistic environments of most English articles. However, their insensitivity to the acceptability condition in the self-paced reading task reveals the late learners' lack of processing ability. Insensitivity to the acceptability condition found in the online reading task can be derived from the learners' processing difficulty and from the failure to activate and retrieve their intact representation in real-time. This observation aligns with the Processing Deficit Approach.

Second, the performance difference between the two tasks can be explained by a task effect. While a grammatical acceptability judgment task captures explicit knowledge, a self-paced reading task captures implicit knowledge (Gass, Behney, & Plonsky, 2013). Thus, it is likely that a different type of knowledge was measured by each type of task. A different performance caused by the type of the task is not new. In their study of the acquisition and use of English articles by Korean learners of English, Kim and Lakshmanan (2008) reported a different performance between an acceptability rating task and a self-paced reading task. They explained that a different field of knowledge was activated in the two tasks, and that the degree of awareness was different across the tasks. Although the reason for the performance difference cannot be conclusively identified by these two explanations, either explanation demonstrates that English articles are processed and understood in a different way by Korean advanced learners of English in a grammatical acceptability judgment task and a self-paced reading task.

7. CONCLUSION

The present study aimed to achieve two goals. The first goal was to examine the role of age of first exposure in the acquisition of English article knowledge by EFL Korean advanced learners of English. Another goal was to investigate the impact of learners' age of first exposure on their online processing of English articles. In order to achieve the research goals, the present study utilized two different types of tasks: a grammatical acceptability judgment task and a self-paced reading task.

The findings from the current study can be summarized as follows. First, in the grammatical acceptability judgment task which tests learners' explicit knowledge on English articles, their age of first exposure did not play a significant role in the acquisition of articles of Types 1 to 4 when their English proficiency reached at a highly advanced level. However, as long as Type 5 articles are concerned, the age of first exposure factor played a role. The late exposure group were less sensitive to the use and semantic features of English articles than the early exposure group. Secondly, age of first exposure was an influential factor for processing articles in the self-paced reading task. The results from the self-paced reading task suggested that articles were processed differently according to learners' age of first exposure. The early age exposure group showed a higher sensitivity to acceptability condition and an overall faster reading times than the late age exposure group. The interaction between the two main factors found in most of the article types further shows the close relationship between the learners' age of first exposure and their article processing. Despite a comparatively weak impact of age of first exposure factor observed in the learners' acceptability judgment task performance, the analyses for reading times data from the self-paced reading task showed that the learners' age of first exposure plays an essential role in processing English articles in real-time.

Given that both participant groups were at a highly advanced level and learned English in EFL situations, the participants' performance of English articles can be attributed to their age of first exposure. Therefore, even in EFL instructional settings where the quality and amount of input is limited, the findings of the study suggest that consistent exposure to English or instruction in early age (minimum 3 hours per week before the age of 12) positively affects learners' ability to process English articles real-time as well as their linguistic representation regarding Type 5 articles. Furthermore, considering that the minimum amount of exposure that the participants received was only 3 hours per week and the target feature is known to be one of the most difficult grammatical items for Korean learners of English, consistently providing sufficient and good quality of input at the L2 learners' early age can be beneficial not only for the learners' acquisition and processing ability of English articles but also for their overall L2 proficiency in the long run. However, starting early in EFL settings where only two 40-minute-long classes per week offered by

public primary schools as compulsory education may not help EFL learners to reach at the level that the early group of the current study achieved. Given that this study yielded positive evidence for starting early, extending 3rd grade primary school English instruction to 3 full hours per week may help EFL Korean learners' long-term success. Such a policy may not impose great burden to both students and teachers while enhancing the effectiveness of early foreign language education.

This study compared the performance of the two different tasks between the early age exposure group and the late age exposure group of highly advanced Korean EFL learners. However, it would be also interesting for future research to examine how the performance of early exposed EFL learners differ from English native speakers' performance, and how each participant's language learning aptitude might affect their performance. In addition, further studies also need to be conducted using other experimental tools such as an eyetracking, an ERP, which will provide further insights on the role of age of exposure in English article acquisition and processing. Despite these remaining issues, the study has a value in that it provided one small, concrete evidence of early start in an EFL setting and a possible direction for future research regarding the important question, "Is it beneficial to start early in an EFL setting?".

Applicable levels: Elementary, secondary, tertiary

REFERENCES

- Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20(3), 242-275.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, England: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Bertram, R., Hyönä, J., & Laine, M. (2000). The role of context in morphological processing: Evidence from Finnish. *Language and Cognitive Processes*, 15(4-5),

367-388.

- Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research*, 13(2), 116-137.
- Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, 2(2), 127-145.
- Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma Press.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44(2), 235-249.
- Butler, Y. G. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition*, 24(3), 451-480.
- Carlson, G., & Sussman, R. (2005). Seemingly indefinite definites. *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, 85, 71-85.
- Celce-Murcia, M., & Larsen-Freeman, D. E. (1999). *The grammar book: An ESL teacher's course* (2nd ed.). Boston, MA: Heinle & Heinle.
- Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, 10(12), 564-570.
- Cummings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369-382.
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-533.
- DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31(3), 413-438.
- Dirix, N., & Duyck, W. (2017). The first- and second-language age of acquisition effect in first- and second-language book reading. *Journal of Memory and Language*, 97, 103-120.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348-368.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41(1), 78-104.
- Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course* (4th ed.). New York, NY: Routledge.
- Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, 51(1), 1-50.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and*

- Cognition*, 27(6), 1411-1423.
- Green, D. W. (2003). The neural basis of the lexicon and the grammar in L2 acquisition. In R. van Hout, A. Hulk, F. Kuiken, & R. J. Towell (Eds.), *The interface between syntax and the lexicon in second language acquisition* (pp. 197-208). New York, NY: John Benjamins.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 50(3), 346-363.
- Huebner, T. (1983). Linguistic systems and linguistic change in an interlanguage. *Studies in Second Language Acquisition*, 6(1), 33-53.
- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3-69.
- Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning*, 60(4), 877-925.
- Ionin, T., Montrul, S., Kim, J.-H., & Philippov, V. (2011). Genericity distinctions and the interpretation of determiners in second language acquisition. *Language Acquisition*, 18(4), 242-280.
- Ionin, T., Montrul, S., & Santos, H. (2011). An experimental investigation of the expression of genericity in English, Spanish and Brazilian Portuguese. *Lingua*, 121(5), 963-985.
- Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118(4), 554-576.
- Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, 24(3), 387-418.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25(4), 603-634.
- Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York, NY: Routledge.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99.
- Kim, K., & Lakshmanan, U. (2008). L2 article semantics and second language processing. In R. Slabakova, J. Rothman, P. Kempchinsky, & E. Gavruseva (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA)* (pp. 103-117). Somerville, MA: Cascadilla Proceedings Project.
- Ko, H., Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic Inquiry*, 41(2), 213-254.
- Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger

- starting age in a minimal input situation. *Second Language Research*, 24(1), 35-63.
- Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, 2(12), 59-67.
- Liu, D., & Gleason, J. L. (2002). Acquisition of the article the by nonnative speakers of English. *Studies in Second Language Acquisition*, 24(1), 1-26.
- Master, P. A. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Master, P. A. (1994). The effect of systematic instruction on learning the English article system. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 99-122). Cambridge, England: Cambridge University Press.
- Master, P. A. (2002). Information structure and English article pedagogy. *System*, 30(3), 331-348.
- 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.
- Meulman, N., Wieling, M., Sprenger, S. A., Stowe, L. A., & Schmid, M. S. (2015). Age effects in L2 grammar processing as revealed by ERPs and how (not) to study them. *PLoS one*, 10(12), e0143328. <https://doi.org/10.1371/journal.pone.0143328/>.
- Muñoz, C. (2010). On how age affects foreign language learning. In A. Psaltou-Joycey & M. Mattheoudakis (Eds.), *Advances in research on language acquisition and teaching: Selected papers from the 14th GALA conference* (pp. 39-49). Thessaloniki, Greece: Greek Applied Linguistics Association.
- Murphy, S. L. (1997). *Knowledge and production of English articles by advanced second language learners*. Unpublished doctoral dissertation, University of Texas at Austin, Austin.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
- Park, T.-S. (2005). English article use by advanced Korean EFL learners. *Foreign Languages Education*, 12(1), 35-59.
- Park, T.-S., & Song, M.-J. (2008). The primary causes of article errors made by Korean advanced learners of English. *English Teaching*, 63(3), 71-90.
- Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37(3), 361-384.
- R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Austria, Vienna. <https://www.R-project.org/>
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16(2), 135-172.
- Snape, N. (2008). Resetting the nominal mapping parameter in L2 English: Definite article

- use and the count–mass distinction. *Bilingualism: Language and Cognition*, 11(1), 63-79.
- Song, M.-J. (2014). Learnability and starting age of English article acquisition: Evidence from CPH. *Language Research*, 50(2), 515-544.
- Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, 65(4), 860-895.
- Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning*, 35(3), 373-403.
- Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10(3), 335-355.
- Trenkic, D., Mirkovic, J., & Altmann, G. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition*, 17(2), 237-257.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4(2), 105-122.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1), 231-270.
- White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 12(3), 233-265.