Artificial Grammar Learning Ability Predicts L2 Processing of English Number Agreement

On-Soon Lee
(Korea University)

Lee, On-Soon. (2016). Artificial Grammar Learning Ability Predicts L2 Processing of English Number Agreement. *Language Research*, 52.1, 87-114.

This study investigates why L2 learners differ in their success in learning their second language, by examining the correlation between learning ability in artificial grammar involving an adjacent dependency and ability in processing English number agreement. Thirty-four Korean L2 learners of English completed two tasks: an artificial grammar learning task and a self-paced reading task. The results showed that (a) the L2 learners are able to track a frequent pattern (e.g., adjacent dependency) occurring in the artificial language; (b) the advanced L2 learners were not sensitive to violations of English number agreement in determiner phrases; and (c) the ability to learn a rule in the artificial grammar learning task was positively correlated with the L2 learners' insensitivity to disagreement in the self-paced reading task. These findings shed light on the effect of the ability to track a rule in an artificial language on L2 sentence processing. The results thus suggest one explanation for why L2 learners' variations appear in language learning, at least in the case of the comprehension of English adjacent number agreement.

Keywords: L2 sentence processing, artificial grammar learning, English number agreement, cognitive ability, individual differences

1. Introduction

Many studies have tested the role of working memory in L2 sentence processing, but there is no general agreement that working memory is significantly correlated with L2 learners' performance (for a review, see Juffs & Harrington, 2011). For example, Juffs (2004, 2005) investigated the relationship between working memory span and online comprehension in both the L1 and the L2 of Chinese, Japanese, and Spanish learners of English. Juffs looked for a correlation between the learners' working memory scores and their mean reading times in the L2 at the critical verb (e.g., *looked*) in temporarily ambiguous sentences (e.g., *After the children cleaned the house looked very neat and tidy*). However, Juffs found no significant correlation between reading times at the critical region and working memory. In more recent work, Havik, Roberts, van Hout, Schreuder, and Haverkort (2009) failed to find an effect of working memory capacity on the L2 processing of Dutch relative clauses. In similar studies, a few Korean researchers have used reading span tests[1)] to examine the effects of working memory on Korean adults' L2 sentence comprehension (J Baik, S-Y Lee, Y Kim, & SJ Lee, 2012; S Baek, 2013; J Baik, S-Y Lee, & Y Kim, 2013; J-H Lee, 2014). Findings showed the effect of working memory in a few types of linguistic structures (e.g., subject-modifying relative clauses or *wh*-questions). The general effect of working memory is still questionable.

In short, the findings from existing L2 studies are not clear enough to conclude that working memory capacity plays an important role in L2 learners' language processing. Moreover, the reading span test popularly employed to measure working memory might be problematic for two reasons: (a) It fails to measure the processing cost of incoming words (for a discussion, see Ariji, Omaki, & Tatsuta, 2003), and (b) it focuses on memorizing the final words of sentences, not on reading sentences (Waters & Caplan, 1996). The test therefore does not provide a clear picture of how cognitive capacity affects L2 learners' performance in sentence processing.

Based on these observations, some recent work has tested the idea that an ability to track a rule from artificial language, which is considered as cognitive capacity, might be a potential predictor of language processing. A few researchers have adopted the artificial grammar learning paradigm

---

1) In the reading span test, an experimenter shows participants a set of sentences written on a sheet of paper and asks them to read the sentences aloud. After reading all of the sentences, they are asked to recall the final word of each. Participants are shown sets with increasing numbers of sentences. The reading span size is defined as the maximum number of sentences for which participants are able to recall the final word.

in attempts to explain why adult monolingual native speakers of English show different performance in online comprehension (Misyak, Christiansen, & Tomblin, 2010a; Misyak & Christiansen, 2012; Farmer, Misyak, & Christiansen, 2012). Taking up the idea that the ability to track a rule in an artificial language is related to the ability to comprehend natural language, the current study addresses why L2 learners achieve different levels of L2 attainment. Specifically, the study investigates the relationship between artificial grammar learning and L2 online comprehension.

2. Background

2.1. Artificial grammar learning (AGL)

It is believed that infants acquire nouns (e.g., *mom* or *dog*) before other words, because nouns are the most frequent words in children's early utterances. One explanation for this phenomenon emphasizes that frequent items are easier to be acquired. With regard to this, Saffran and her colleagues proposed that sensitivity to frequent patterns occurring in human languages (e.g., frequent phoneme sequences or frequent structure) is related to a mechanism for acquiring language. To measure such sensitivity, they invented an artificial language (e.g., a novel language consisting of English phonemes) in which they manipulated the frequency of patterns of phoneme sequences. In an earlier study, Saffran, Aslin, and Newport (1996) investigated whether infants are sensitive to frequent phoneme sequences in an auditory string of speech without any linguistic cues (e.g., pauses). They were specifically interested in whether infants are able to use transitional probability to identify words in a continuous stream of speech. For example, the English syllable *pre* precedes a small set of syllables, including *ty*, *tend*, and *cedes.* In the stream of speech addressed to babies, the probability that *pre* is followed by *ty* is relatively high. However, the probability that *ty* is followed by *ba*, as in *pretty baby*, is extremely low; this is called transitional probability. This difference in transitional probability is a clue that *pretty* is a word and *tyba* is not. The ability to track transitional probabilities may thus be useful to infant learners in discerning words. The study's findings suggested that an ability — in this case, learning the patterns of certain syllables occurring together more frequently than others — plays a role in language learning. That is, if people can track phoneme sequences or linguistic structures frequently occurring in language, such as a transitional pattern (i.e. statistical pattern), most linguistic properties in actual human language might be learnable in this way.

Since the work of Saffran and her colleagues, researchers have extensively adopted the artificial language paradigm in order to examine the role of the ability to track frequent patterns in learning grammatical rules, such as rules of phrase structure (Thompson & Newport, 2007; Takahashi & Lidz, 2008) and syntax-like regularities (Gómez & Gerken, 1999). In the earlier research, Gómez and Gerken (1999) investigated whether 12-month-old infants could learn a very simple grammar rule manipulated in an artificial language as in (1a). In this artificial language, a pair between the first (*vot*) and the last (*jic*) nonwords in (1a) forms a grammatical nonadjacent dependency (*vot-jic*) with three intervening nonwords (*pel pel pel*). However, the opposite pattern *(* jic-vot*) in (1b) is not grammatical.

(1) a. **vot** pel pel pel **jic**
  b. ***jic** pel pel pel **vot**

In their study, the experiment consisted of training and test sessions. In the training session, the infants were exposed to a variety of auditory grammatical strings, such as that in (1a), for two minutes. In the following test session, they judged whether each string was grammatical, using the head-turn preferences procedure: 10 grammatical, as in (1a), and 10 ungrammatical, as in (1b). The prediction was that infants would listen longer to grammatical strings than to ungrammatical strings, indicating acceptance and rejection, respectively. The infants successfully distinguished grammatical strings from ungrammatical strings, demonstrating that infants are capable of abstracting a simple structure from an artificial language they have never heard before.

To summarize, the findings from previous research on artificial language suggest that language learners are very sensitive to frequent patterns at the level of words (Saffran et al., 1996) and the level of structures (Gómez & Gerken, 1999, 2000; Gómez, 2002; Gómez & Maye, 2005). Furthermore, many researchers believe that learners' ability to abstract a rule from an artificial language is associated with their ability to comprehend natural language; that is, the same learning mechanism applies to both artificial grammar learning and natural language acquisition (Gómez & Gerken, 1999, 2000). However, this conclusion is far from certain. Even though previous work has successfully tested learners' ability to abstract a rule from input, the theoretical assumption of a correlation between the ability to track a rule from artificial language and processing ability in natural language has rarely been tested, particularly in L2 acquisition.

2.2. Validity of the AGL paradigm in L1 and L2

As pointed out in the previous section, few studies have empirically tested the idea that the ability measured by artificial grammar learning is similar to the mechanism of language learning. Among the little research to test this possibility is that conducted by Misyak and her colleagues (Misyak et al., 2010a; Misyak & Christiansen, 2012). They hypothesized that the way to abstract a rule from artificial language, as in (2), is related to the way to process English subject or object relative clauses (SRCs and ORCs), as in (4), due to the similarity of the dependency formed by the intervening words. For example, the intervening words between the head noun (the filler, *reporter*) and its gap marked with an underscore as in (4) create a nonadjacent dependency. In (2), a pair of the first *(pel, dak, vot*) and last (*rud, jic, tood*) words forms three pairs of nonadjacent dependences (*pel-rud, dak-jic, vot-tood*). That is, the dependencies in artificial grammar resemble those found in relative clauses in a limited but relevant way: they hold across intervening material. The examples in (3) are ungrammatical according the rule as in (2).

(2) a. **pel** wadim **rud**
    b. **dak** kicey **jic**
    c. **vot** loga **tood**

(3) a. ***pel** wadim **jic**
    b. ***dak** kicey **tood**
    c. ***vot** loga **rud**

(4) a. SRC: The reporteri that ___i attacked the senator admitted the error.
    b. ORC: The reporterj that the senator attacked _____ j admitted the error.

With these materials, Misyak and Christiansen (2012) hypothesized that participants would be able to track such rule in this artificial language if they were sensitive to the frequency of patterns occurring in an artificial language. In their study, twenty monolingual native English speakers completed two experiments in one day: an artificial grammar learning task and a self-paced reading task. In the first experiment, designed to test their sensitivity, participants first were instructed to listen to 432 grammatical strings consisting of three nonsense words, as in (2), in a training session. In the following test session, each participant read and judged the grammaticality of 12 strings, six grammatical, as in (2), and six ungrammatical, as in (3).

In the second experiment, participants read 40 English sentences with relative clauses, as in (4). The sentences were presented in a word-by-word, moving window, self-paced reading task (Just, Carpenter, & Woolley, 1982). The participants saw sentences with 20 subject relative clauses and 20 direct object relative clauses along with 48 filler sentences. A yes/no comprehension probe followed each sentence. Participants also completed a reading span task (Waters & Caplan, 1996) intended to test the effect of working memory.

The results from the experiments showed that (a) participants were able to abstract a rule from artificial language (their overall judgment accuracy rate was 62.1%); (b) learners' performance in the artificial grammar learning task was positively correlated with their performance in processing English relative clauses (r = .38, p < .05); and (c) there was no correlation between scores on the reading span task and learners' performance in processing English relative clauses.

A follow-up study (Misyak & Christiansen, 2012) conducted another experiment with the same participants, who completed two tasks: an artificial grammar learning task including adjacent dependencies and a phonological typicality task.2) The results again confirmed the validity of artificial grammar learning ability to predict variations in L1 language performance. However, there was an issue with the materials the study used. The patterns of the phonological typicality task, which was used for language comprehension, did not correspond to the rules manipulated in the artificial grammar learning task. The present study, therefore, employs more appropriate materials. In spite of the problem with materials in the later study, overall, the research done by Misyak and colleagues suggests the usefulness of the artificial grammar paradigm for investigating the relationship between cognitive capacity and language learning.

In the same vein, Robinson (2005) used a task from two previous artificial grammar studies on L1 acquisition (Reber, Walkenfeld, & Hernstadt, 1991; Knowlton & Squire, 1996) as a measurement of L2 learners' implicit learning ability. The task tested the ability to abstract a rule from an artificial grammar without explicit instructions. Robinson also extended the previous studies by examining the role of implicit learning ability as measured by artificial grammar learning in learning a novel natural language (i.e., Samoan). Robinson found that (a) the L2 learner successfully abstracted a rule in an artificial grammar learning task, as the L1 speakers did in Reber et al.'s and Knowlton and Squire's studies; and (b) their success in implicit learning did not help the L2 learners to acquire natural language (i.e., Samoan). Robinson's findings alone, however, are

---

2) The following examples are from this task (Misyak & Christiansen, 2012, p. 313).
(i) Noun-like homonym with noun/verb resolution
Chris and Ben are glad that the bird perches [seem easy to install]/[comfortably in the cage].
(ii) Verb-like homonym with noun/verb resolution
The teacher told the principal that the student needs [were not being met]/[to be more focused].

not sufficient to understand how relevant implicit artificial grammar learning might or might not be in L2 learning. To further explore this question, the current study examines how relevant artificial grammar learning as a cognitive ability is in L2 learners' processing of English number agreement.

2.3. Previous studies on L2 processing of English number agreement

L2 learners typically have trouble correctly using the English plural *-s* in production and even in online comprehension. For example, several studies have shown L2 learners to be insensitive to the erroneous usage of the morpheme in comprehension during real-time processing (Jiang, 2004, 2007; Chen, Shu, Liu, Zhao, & Li, 2007; S-W Kim, 2011; S-W Kim, M-K Park, & W Chung, 2012). Nevertheless, other studies have shown findings that suggest the opposite (Sato & Felser, 2008; Wen, Miyao, Takeda, Chu, & Schwartz, 2010).

In the earlier research, Jiang (2004) investigated whether advanced Chinese L2 learners could detect the erroneous usage of English plural *-s* in a self-paced reading task. He employed materials using sentences like those in (5) and (6).

(5) a. The bridges to the island *were* about ten miles away.
    b. * The bridge to the island *were* about ten miles away.

(6) a. the visitor took *several* of the rare coin*s* in the cabinet.
    b. * The visitor took *several* of the rare coin in the cabinet.

Jiang assumed that if the learners had knowledge of plural *-s*, they should be able to detect the disagreements as in (5b)/(6b) in real-time processing, as English native speakers do. If the participants recognized the ungrammaticality, the reading time (RT) at a critical region such as the verb *were* in (5b) would be significantly longer than at the critical region in a grammatical sentence, such as the verb *were* in (5a). The findings showed that the RTs of the English native speakers but not the L2 learners were significantly longer at the critical region in sentences like (5b) than in those like (5a). Jiang (2007) conducted a second, similar study with the same results, confirming L2 learners' insensitivity to violations of English number agreement.

On the other hand, Sato and Felser (2008) conducted a speeded-grammaticality-judgment task in order to examine whether L2 learners can recognize violations of subject-verb (dis)agreement, as in (7).

(7) a. *She* rarely flirt*s*.
 b. **She* rarely flirt.

Participants were asked to read sentences presented word-by-word at the center of a computer screen, and then, at the end of each sentence, to decide whether it was well formed. The findings showed that all the L2 groups' judgment-accuracy rates were well above chance (German: 93.4%; Japanese: 83.8%; Chinese: 62.8%). This indicates that the L2 learners were able to use knowledge of the English plural *-s* in subject-verb agreement. Along the same lines, Wen et al. (2010) investigated the effect of L2 proficiency on detecting English number (dis)agreement with Chinese and Japanese advanced L2 English learners. They employed a self-paced reading task with sentences such as those in (8).

(8) a. Jill sold *this* beautiful *house* to her niece every evening.
 b. *Jill sold *this* beautiful *houses* to her niece every evening.
 c. Jill sold *these* beautiful *houses* to her niece every evening.
 d. *Jill sold *these* beautiful *house* to her niece every evening.

(8b) and (8d) show violations of English number agreement between demonstratives (e.g., *this/these*) and head nouns (e.g., *house/houses*). If the L2 learners were sensitive to such disagreement, they should read significantly more slowly at the critical regions, such as *house/houses*, in the ungrammatical sentences than at the same regions in the grammatical sentences (8a)/(8c), as English native speakers do. The study's results showed that the RTs from the advanced L2 learners, but not from the intermediate L2 learners, were indeed significantly longer at the critical regions in the ungrammatical sentences. This indicates that L2 learners are able to acquire and use the English plural *-s*, and it also suggests that L2 learners' proficiency does matter in online comprehension.

A few studies on the processing of English number agreement by Korean adult learners of English have provided evidence that Korean L2 learners have difficulty recognizing violations of English subject-verb (dis)agreement (S-W Kim, 2011). Kim and his colleagues replicated Chen, Shu, Liu, Zhao, and Li's (2007) study in order to examine whether intermediate Korean L2 learners have knowledge of the English plural *-s* in subject-verb (dis)agreement. Kim's (2011) study employed a self-paced reading task with experimental stimuli as in (9). The 32 native Korean-speaking adults who participated were asked to read sentences presented word-by-word; at the end of each sentence, they were asked to decide if it was grammatical.

(9) a. The price of the car *was* too high.
 b. The price of the car*s* *was* too high.
 c. *The price of the car*s* *were* too high.
 d. *The price of the car *were* too high.

The results showed that the L2 learners' judgment accuracy rate was above chance (68.4%). The participants' judgment accuracy rate was 82.5% for sentences like (9a)/(9b) and 54.5% for sentences like (9c)/(9d). This indicates that Korean adult L2 learners might lack performance capacity in real-time processing. This means that Korean L2 learners showed the difficulty of detecting ungrammaticality of English number agreement. Findings from both studies therefore confirmed that Korean L2 learners are not able to discern (dis)agreement in sentences with English plural *-s*. L2 learners' sensitivity to ill-formed English agreement is poor in real-time processing. It is possible that cognitive ability might explain Korean adult L2 learners' poor performance in online comprehension.

To summarize, the previous studies have shown that some advanced L2 learners whose first language (L1) does not have number agreement are sensitive to violations of English number agreement with the morpheme *-s* in real-time processing, while others are insensitive to the same violations. To explain such variations in L2 learners' attainment, the present study examined the hypothesis that the better the L2 learners' cognitive ability (i.e., the ability to learn a rule from an artificial language), the better they will detect (dis)agreement in English. More specifically, this study also investigates what causes variations in Korean adult L2 learners' performance in processing English number agreement.

2.4. Research questions

Misyak and Christiansen (2012) showed empirically that artificial grammar learning is connected with language acquisition. Their study reported that the participants were able to discern the grammaticality of artificial language at a fairly high rate (62.1%, *SD* = 14.3). However, only English native speakers participated in their experiments. The present study seeks to extend Misyak and Christiansen's study by exploring the relationship they found between artificial grammar learning and natural language learning in an L2 context.

In addition, Misyak and Christiansen's study used some problematic materials, in which the sentences in the language comprehension did not correspond to the rules manipulated in the artificial grammar learning task. Moreover, the findings of other previous studies (Jiang, 2004, 2007; Chen et al., 2007; W-S Kim, 2011; Kim et al., 2012), showing L2 learners' insensitivity to English plural *-s* during real-time processing, might be due to factors (e.g., a cognitive ability measured by artificial grammar learning) other than their knowledge of English number agreement. Therefore, the current study's materials test L2 learners' sensitivity to English number agreement between a demonstrative and a noun, which corresponds to an adjacent dependency pattern in artificial grammar; the predictive relation between a demonstrative and a noun is a case of adjacent dependency (e.g., *this doll, *this dolls, these dolls, *these doll*). For example, *this* predicts the following word *doll* in DPs.

Finally, this study explores the relation between learning an adjacent rule from artificial language and processing English adjacent number agreement in DPs. The logic of this question is that the way to learn adjacent dependency pattern in artificial grammar is similar to the way to process English adjacent number agreement, due to the similarity of the adjacent dependencies formed in both artificial grammar learning task and a self-paced reading task. This question further contributes to research on whether the learners' ability to implicitly track a rule in artificial grammar might be relevant to processing natural language.

The study addresses three research questions. The first two questions might be necessary to understand the relationships between AGL task and L2 sentence processing as a preliminary step so as to better understand the third question. Therefore, the first two will provide results that can be compared with those of previous research using AGL tasks and sentence processing, respectively. The third is the most interesting question of this study, as it not only builds on but extends the previous research.

1. Do Korean adult L2 learners of English implicitly track/learn adjacent dependencies in an artificial grammar learning task in the same way that English native speakers do? Are they sensitive to the frequent pattern in artificial language?
2. Are L2 learners able to detect the ungrammaticality of English number agreement in DPs?
3. Does the ability to learn a rule in an artificial grammar predict learners' performance in processing English number agreement?

3. Method

3.1. Participants

Thirty-four Korean learners of English (28 women and 6 men; mean age = 20.9, *SD* = 1.3, range = 19-25), all undergraduates, participated. Each participant completed the experiment in a single, individual session. Participants received course credit or monetary compensation. To assess the effect of the L2 learners' English proficiency on their sentence-processing performance, a C-test adapted from one used by Schulz (2006) was conducted, with a 10 minute time limit. The C-test had 40 items, each consisting of the first half of a word plus a blank (see Appendix A). The highest possible score was 40. Participants with scores from 27 to 40 were classified as advanced, and those with scores from 16 to 27 were classified as intermediate.[3] Table 1 shows the language profiles of all participants.

Table 1. Participants' Language Background and English Proficiency

	Male: Female	Mean Age	C-test Scores
Groups			
Advanced	3:18 (N = 21)	20.8	31.1 (SD = 2.04)
Intermediate	3:10 (N = 13)	21	22.8 (SD = 4.07)

An independent *t*-test showed that the advanced L2 learners' English proficiency was relatively higher than that of the intermediate L2 learners according to their C-test scores [t (32)= 7.853, p < .05].

3.2. Materials

3.2.1. The artificial grammar learning task

The artificial grammar was adapted with modification from Friederici, Steinhauer, and Pfeifer's (2002) materials. In this grammar, adjacent dependencies occur within phrases as in (10).

(10) a. $S_{entence} \rightarrow AP_{hrase}\ BP_{hrase}$
b. $AP_{hrase} \rightarrow DA$
c. $BP_{hrase} \rightarrow B\ AP_{hrase}$

Within phrases, a D constituent always predicts and occurs prior to an A constituent as in (10b). According a B phrase (BP) in (10c), an A phrase (AP) follows every B. Thus, a sentence consists of two phrases, AP and BP in (10a); AP always precedes BP in a sentence; every B in

3) The score 27, the cutoff value, might be questionable. However, this does not matter for the purpose of this study because this study needs proficiency measure according to which this study evaluates participant's proficiency level relative to that of the other participant's proficiency level (Schultz, 2006, p.76).

BP is followed by AP; and D occurs prior to A in AP. In this rule, the predictive-based learning is necessary because D is followed by A in (10b). According to this rule, the language was instantiated by nine monosyllabic nonsense words (*jux, dupp, hep, lum, meep, rauk, sig, tam, zoet*) belonging to three lexical categories; there are four A members, four B members, and one D member. Sixty-four grammatical sentences were generated according to the adjacent dependency rules listed in (10). Ungrammatical sentences were created by replacing one nonword with another taken from a different category. For example, suppose that a grammatical sentence contained the following sequence of category constituents: D A B D A. Its ungrammatical counterpart could be created by replacing the second D with an A: *D A B A A (e.g., *jux dupp tam jux dupp* vs. **jux dupp tam hep dupp*).

With these sentences, the artificial grammar learning task was presented using a word-by-word paradigm on a computer running E-Prime (Version 2.0), as in a self-paced reading task.[4] Figure 1 illustrates how the task was organized.

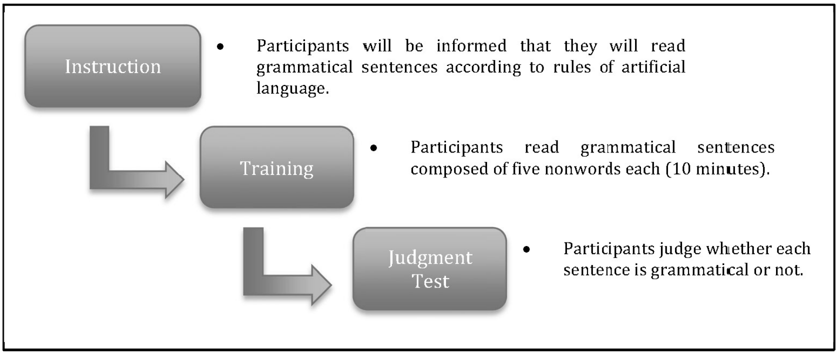

Figure 1. Procedure of the artificial grammar learning task.

As in Figure 1, artificial grammar learning tasks consist of a training and a test session. Before starting the task, participants were informed that they would read a series of sentences from an artificial language,

4) The reason to use a word-by-word moving window in both of tasks is that readers naturally read the sentence in a left-to-right, a word-by-word fashion.

all of which were possible sentences including five nonwords according to the rules of the artificial grammar (see the example in [10]). In the training session, each sentence was presented automatically, after a fixation point appeared on the screen for 250 milliseconds. Next, each word of the given sentence appeared automatically for 250 milliseconds (ms) on the left side of the computer screen, and then the whole sentence consisting of these five words was held for one second on the computer screen. Other sentences followed, in a random order, with the same presentation method. The training session lasted approximately 10 minutes.

In the test session, each sentence was presented one word at a time on the computer screen, left to right, in a cumulative, moving-window format as a participant pushed a space bar (Just et al., 1982). Participants were asked to give their judgment about the grammaticality of 40 sentences, half of which were grammatical, and half of which were ungrammatical. Half of the grammatical sentences were identical to sentences in the training session in order to test to what extent learners either memorized the sentences or learned the rules of the artificial language. The other half of the grammatical sentences were new; that is, they were structurally similar but not identical to those in the training session.5) This division into identical versus structurally similar but not identical sentences was important to test whether the learners were capable of tracking the rule (e.g., Misyak & Christiansen, 2012).

3.2.2. The self-paced reading task

The self-paced reading task's stimuli consisted of 24 sets of two sentences each, in four conditions as illustrated in Table 2. Each sentence contained nine regions. The *This*-Nsg and *These*-Npl conditions are grammatical, because the demonstrative and the noun match in number. The *This*-Npl and *These*-Nsg conditions are ungrammatical because the demonstrative and the noun mismatch in number. The 24 sentences were distributed in a Latin square design across four lists, randomly assigned to participants

---

5) Results show that the accuracy rate for new grammatical sentences is not significantly different from that of the identical grammatical sentences were (77% vs. 85%, p = 0.52). This indicates that L2 learners appropriately learn phrase structure rules in artificial grammar learning.

so that they each saw only one condition of each experimental item. Each list also contained 48 fillers. All fillers were simple sentences: 24 grammatical sentences (e.g., *Fred chased the rabbits and gave cookies to them*) and 24 ungrammatical sentences containing erroneous past tense forms of verbs (e.g., **Jessica left the theater while the ballerina tooked breaks*). In total, each participant read 72 sentences. After reading each sentence, participants were asked whether the sentence was grammatical. The ratio of grammatical to ungrammatical items was balanced.

Table 2. Sample Set of Experimental Stimuli

Conditions	Regions								
	1	2	3	4	5	6	7	8	9
This-N_{sg}	Jessie	gave	this	***doll***	to	her	daughter	last	year.
This-N_{pl}	*Jessie	gave	this	***dolls***	to	her	daughter	last	year.
These-N_{pl}	Jessie	gave	these	***dolls***	to	her	daughter	last	year.
These-N_{sg}	*Jessie	gave	these	***doll***	to	her	daughter	last	year.

3.2.3. Procedure

The participants first completed the artificial grammar learning task and then the self-paced reading task. The C-test and a language background questionnaire followed (Appendixes A and B). Each individual session lasted approximately 50 minutes.

4. Results

4.1. AGL, L2 processing of English number agreement, and proficiency

This section reports the results from the three tasks: the artificial grammar learning task, the self-paced reading task, and the proficiency test (i.e., C-test). For the artificial grammar learning task, the analysis was conducted on the judgment accuracy rates as a dependent measure. For the self-paced reading task, the analysis was conducted on the grammaticality judgment accuracy rates as a dependent measure. Although the reading times were recorded, measures of interest for the analyses were grammaticality judgment scores that served as offline language ability. Because AGL task also measures a metacognitive component for participants' judgments, offline language comprehension scores were suitable for comparisons across the two tasks (Misyak & Christiansen, 2012, p.312). Moreover, grammaticality judgment scores are appropriate to show variations in participants' performance.

Table 3 presents the descriptive statistics. The results of each of the tasks will be discussed separately in terms of the findings from previous studies.

Table 3. Descriptive Statistics of the L2 Learners' Scores on the Three Tasks

Task	Dependent measures	Mean (*SD*)	Observed range (%)	Possible range (%)
AGL	Number correct (40)	81.44 (15.8)	0-100	1-100
Self-paced reading task	Number correct (24)	66.32 (18.9)	30-100	1-100
Proficiency test	Number correct (40)	27.94 (5.01)	40-92.5	1-100

Note. AGL = artificial grammar learning

As shown in Table 3, the average of the participants' judgment accuracy rates in the artificial grammar learning task was 81.44% (*SD* = 15.8), similar to the rates of the English native speakers (69.2%, *SD* = 24.7) in Misyak and Christiansen's (2012) study. Thus, the current study replicated their finding, showing that L2 English learners are able to abstract this rule (i.e., adjacent dependency) from an artificial grammar.

Only a few prior studies on artificial grammar learning have argued that the rule-tracking ability in an artificial grammar is relevant for acquiring language at the syntactic level (Gómez & Gerken, 1999, 2000; Gómez, 2002; Gómez & Maye, 2005). Moreover, those studies have tested only native speakers of English (for a summary, see Misyak & Christiansen, 2012). The current study's findings thus contribute to the literature on artificial grammar learning by showing that native Korean-speaking L2

English learners are capable of implicitly tracking adjacency patterns in an artificial language. These results are consistent with the findings of previous studies in English native group (Knowlton & Squire, 1996; Saffran et al., 1996; Gómez, 2002; Misyak & Christiansen, 2012). To recap, L2 learners are sensitive to the frequent pattern occurring in a novel language (i.e., artificial language). This might raise the possibility to examine the effect of artificial grammar learning ability in L2 context.

Let us turn to the results of the self-paced reading task. The L2 learners' average judgment accuracy rate was 66.32% (*SD* = 18.9) for target items, and 83% (*SD* = 6.7) for fillers. Broken down by L2 proficiency, the advanced group's average rate was 71% (*SD* = 19.6) for target items, and 86% (*SD* = 6.9) for fillers, while the intermediate group's average rate was 58% (*SD* =15.7) for target items and 79% (*SD* = 3.3) for fillers. The findings are consistent with findings reported by previous studies (Kim, 2011; Kim et al., 2012). For example, in Kim's (2011) study, the overall judgment accuracy rate from 32 Korean L2 learners of English was 68.4%: 82.5% for grammatical sentences and 54.5% for ungrammatical sentences. In the current study, in the case of the advanced L2 learners, their accuracy rate for grammatical sentences was 91.5%, whereas their accuracy rate for ungrammatical sentences was 51%. In short, these studies show that advanced L2 learners are not sensitive to English number disagreement with the plural morpheme *-s* in real time processing.

Again, the descriptive results from the two separate tasks are consistent with the major finding of previous studies, indicating that L2 learners are sensitive to frequent patterns occurring in artificial language. Advanced L2 learners also had difficulties to discern number disagreement in real-time sentence processing. These are preliminary results to better understand the following section.

4.2. Relationship between artificial grammar learning and self-paced reading tasks

The most important objective of this study was to look at the correlation between the ability to learn a rule in an artificial grammar and online comprehension. Table 4 shows that the judgment accuracy rate from the artificial grammar learning task is positively correlated with the judgment accuracy rate from the self-paced reading task (r = .41, p < .05); that is, the L2 learners who were better at learning the rule in the artificial language were also better at detecting (dis)agreement in English.

Table 4. Correlations Between Artificial Grammar Learning and Self-Paced Reading Tasks

Task	AGL	Self-paced reading	Proficiency test
AGL	1		
Self-paced reading	.41*	1	
Proficiency test	.50**	.39*	1

Note. AGL = artificial grammar learning; * p < .05; ** p < .01

To determine how well the grammaticality accuracy rate in the artificial grammar learning task predicts L2 comprehension, a simple linear regression was conducted; both variables were normally distributed, so Pearson's correlation coefficient was calculated. It found a moderate correlation between the judgment rates of artificial grammar learning and grammaticality judgment rates of English number agreement. The regression coefficients of the simple linear regression showed that mean grammaticality accuracy in AGL task predicts the grammaticality judgment accuracy of English number agreement (*ßeta* = .41, p < .05). This means that the participants' mean grammaticality accuracy in the AGL task accounts for 17% of their English grammaticality judgment accuracy, which is significant at the level of p < .05. Moreover, Table 4 shows a considerable correlation between L2 learners' proficiency level and their performance in real-time processing (Sagarra & Herschensohn, 2010; Wen et al., 2010). The mean score from the proficiency test is positively correlated with the judgment accuracy rates in the artificial grammar learning task (r = .50, p < .01) and the self-paced reading task (r = .39, p < .05). This indicates that the advanced L2 learners did better than the intermediate L2 learners in tracking a rule in an artificial language. In addition, the advanced L2 learners were relatively sensitive to (dis)agreement than the intermediate learners.

The most important finding is that the L2 learners who were better at implicitly tracking a rule (i.e., adjacent dependency) in the artificial language were more sensitive to ungrammaticality in English number agreement. Why does this happen? Both of AGL task and a self-paced reading task are similar in important respects. First, the way to process adjacent dependency pattern occurring in artificial grammar is similar to the way to process adjacent number agreement in a self-paced reading task. In the case of artificial grammar learning, the one (i.e., D in A phrase) is followed by the other (i.e., A in A phrase), as *this* in a DP predicts *doll* in a self-paced reading task. This predictive-based learning in AGL task might be related with L2 processing of English number agreement between demonstratives and head nouns (e.g., Misyak, Christiansen & Tomblin, 2010b). Secondly, both require attention (Hisao & Reber, 1998; Schmidt, 2001). During the processing of incoming information, attention "is centered on the on-line processing of the sensory data (Perruchet & Vintner, 2002, p.45). In both tasks, attention is necessary to track the frequent pattern (i.e., chunk) in artificial language and to discern number disagreement in natural language. The ability to track an adjacent dependency in an artificial language, therefore, might predict L2 learners' performance in processing English adjacent number agreement.

This differs from Robinson's (2005) finding that learning artificial grammar did not help L2 learners learn Samoan. This study reports the first finding of a positive relationship between artificial grammar learning ability and natural language acquisition by L2 learners. Two possible reasons can explain this. The AGL task used in his study differs from that used in the current study. While the current AGL task includes the rule of phrase structures, the AGL task in Robinson's (2005) experiment only manipulates the frequencies of chunk occurring together in artificial language. The other reason is that there is no similar mechanism (e.g., adjacent dependency) between learning AGL and learning Samoan. However, the current experiment suggests that the way to learn a rule in AGL task is similar to the way to process an adjacent number agreement in natural language due to the similarity of adjacent dependency in both tasks.

5. Discussions

This study addressed three research questions. First, it attempted to examine whether L2 learners are capable of implicitly learning an adjacency pattern from artificial language to generalize the finding of Misyak and Christiansen (2012). The results provided empirical evidence that L2 learners can track an adjacency relationship in an artificial grammar. In other words, L2 learners are sensitive the frequent pattern so that they can learn a rule of language in novel language, as proven in previous studies (Saffran et al., 1996; Gómez & Gerken, 1999, 2000; Misyak & Christiansen, 2012).

The second research question looked at whether advanced L2 learners are able to recognize violations of number agreement in sentences with English plural *-s* in DPs. However, results from the self-paced reading task showed that advanced Korean L2 learners, along with intermediate learners, are not sensitive to disagreement between demonstratives (e.g., *this/these*) and the following nouns (e.g., *doll/dolls*). This is consistent with findings from previous studies (Kim, 2011; Kim et al., 2012): L2 learners' performance is poor in detecting errors of English number agreement. One explanation for this emphasizes that the variation in cognitive ability might be correlated with L2 learners' performance in real-time processing.

Finally, this study explored the potential relationship between the ability to learn an adjacency pattern in artificial language and the performance in processing English adjacent number agreement. The results showed that the ability to learn an artificial grammar rule influenced L2 learners' judgment accuracy in processing English number agreement. This implies that the degree of the ability to learn an artificial grammar is likely to explain individual differences in Korean adult L2 learners' performance.

Findings from this study provide implications regarding the cognitive psychology of L2 acquisition and L2 pedagogy. First, the ability to track a rule from artificial language is potentially a predictor of degree in learners' learning success, in the case of English adjacent number agreement. Even though several studies have tried to explain individual variation in L2 processing by focusing on the role of working memory, they have not been able to show a consistent effect (for reviews, see Juffs & Harrington, 2011; Juffs & Rodríguez, 2015). Alternatively, the ability to track a rule from artificial language is considered as a kind of cognitive ability, allowing a learner to be measured on the basis of memory as well as the ability to process subsequent input. The study therefore provides additional evidence to explain individual differences in the sense that variation in cognitive ability measured via the artificial grammar learning task affected the ability to process English number agreement. Second, the effect of the frequency of the rules or patterns manipulated in AGL tasks is important for L2 pedagogy in the classroom. Ellis (2002) argued that a frequency effect and chunking are powerful influences on L2 learning. If so, in naturalistic exposure to the L2, chunk-strength in the input and its manipulation should facilitate L2 learning. On this view, findings from this study support Ellis's position; in this study, L2 learners successfully learned an adjacency rule frequently occurring in artificial language. This would imply that manipulations of chunking and frequency in the input are able to facilitate L2 grammar learning in the classroom, in particular, lexical learning (Ellis, 2002).

However, there are limitations in the current study. As Robinson (2005) pointed out, learning a natural language requires form-meaning mapping; however, no such mapping is needed in the case of artificial grammar learning, in which the sentences have no meaning. In addition, this study did not conduct a reading span task, which would have allowed a clear comparison between working memory capacity and rule-tracking ability in artificial grammar learning. Despite these limitations, the findings reported here are not spurious; however, further research should be conducted with an improved experimental design.

6. Conclusions

This paper has mainly reported the effect of the ability to implicitly track a rule in an artificial grammar on L2 learners' processing of English number agreement. A general effect of the ability to abstract a rule appeared in L2 learners' processing of English number agreement. The results showed that L2 learners who did well in the artificial grammar learning also did well in detecting (dis)agreement between demonstratives and nouns, whereas L2 learners who were poor at learning the rule in the artificial grammar also had lower judgment accuracy in English. These findings from the L2 group suggest that artificial grammar learning ability can help to explain differences in sentence processing ability. Moreover, the results raise the possibility that rule-tracking ability might be the long-sought component of working memory capacity that could help explain differences in L2 learning ability. Of course, artificial grammar learning ability is not the only factor explaining why some L2 learners learn more quickly and more successfully than others, but its role in successful L2 processing and acquisition is worthy of further consideration. Moreover, further studies with different types of linguistic phenomena (e.g., island effects, nonlocal agreement) are necessary to expand the better understanding of artificial grammar learning ability.

References

Ariji, Kenji, Omaki, Akira, and Tatsuta, Nano. (2003). Working memory restricts the use of semantic information in ambiguity resolution. In P. Slezak (Ed.), *Proceedings of the 4th International Conference on Cognitive Science* (pp. 19-25). Sydney, Australia: University of New South Wales.

Baek, Soondo. (2013). Working memory as an independent predictor of second language sentence comprehension. *Korean Journal of Applied Linguistics* 29.4, 1-25.

Baik, Juno, Lee, Sun-Young, and Kim, Youngjoo. (2013). Working memory: Correlation with second language development and its measuring methodology. *Korean Journal of Applied Linguistics* 29.1, 127-150.

Baik, Juno, Lee, Sun-Young, Kim, Youngjoo, and Lee, Sun Jin. (2012). Working memory and grammaticality judgment test. *Korean Journal of Applied Linguistics* 28.1, 31-58.

Chen, Lang, Shu, Hua, Liu, Youyi, Zhao, Jingjing, and Li, Ping. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism: Language and Cognition* 10.2, 161-174.

Ellis, Nick. (2002). Frequency effect on language acquisition. *Studies on Second Language Acquisition* 24.2, 143-188.

Farmer, Thomas, Misyak, Jennifer B., and Christiansen, Morten. H. (2012). Individual differences in sentence processing. In M. Spivey, M. Joannisse, & K. McRae (Eds.), *Cambridge handbook of psycholinguistics* (pp. 353-364). Cambridge, UK: Cambridge University Press.

Friederici, Angela, Steinhauer, Karsten, and Pfeifer, Erdmut. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences* 99, 529-534.

Gómez, R. L. (2002). Variability and detection of variant structure. *Psychological Science* 13, 431-436.

Gómez, Rebecca L. and Gerken, LouAnn. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109-135.

Gómez, Rebecca L. and Gerken, LouAnn. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences* 4, 178-186.

Gómez, Rebecca L. and Maye, Jessica. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy* 7, 183-206.

Havik, Else, Roberts, Leah, van Hout, Roeland, Schreuder, Robert, and Haverkort, Marco. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning* 59, 73-112.

Hsiao, Andrew and Reber, Arthur. S. (1998). The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious. In M. Stadler & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 495-531). Thousand Oaks, CA:Sage.

Jiang, Nan. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics* 25, 603-634.

Jiang, Nan. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning* 57, 1-33.

Juffs, Alan. (2004). Representation, processing, and working memory in a second language. *Transactions of the Philological Society* 102, 199-225.

Juffs, Alan. (2005). The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research* 21, 121-151.

Juffs, Alan and Harrington, Michael W. (2011). Aspects of working memory in L2 learning. *Language Teaching: Reviews and Studies* 42, 137-166.

Juffs, Alan and Rodríguez, Guillermo A. (2015). *Second language sentence processing*. New York: Routledge.

Just, Marcel A., Carpenter, Patricia A., and Woolley, Jacquelin. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental*

Psychology: General 111, 228-238.

Kim, Sun-Woong. (2011). An integrated analysis of subject-verb agreement: Mismatch in English. *Studies in Generative Grammar* 21.3, 459-483.

Kim, Sun-Woong, Park, Myung-Kwan, and Chung, Wonil. (2012). On the relationship between linear proximity and subject verb mismatch in Korean L2 learners of English: An ERP study. *Studies in Generative Grammar* 22.3, 529-556.

Knowlton, Barbara and Squire, Larry. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory and Cognition* 22, 169-181.

Lee, Jin-Hwa. (2014). Working memory and a subject-object asymmetry in L2 processing of English long-distance wh-questions. *Korean Journal of Applied Linguistics* 30.4, 115-139.

Misyak, Jennifer B. and Christiansen, Morten H. (2012). Statistical learning and languages: An individual difference study. *Language Learning* 62, 302-331.

Misyak, Jennifer B., Christiansen, Morten H., and Tomblin, Bruce. (2010a). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, 1:31. doi: 10.3389/fpsyg.2010.00031.

Misyak, Jennifer B., Christiansesn, Morten H., and Tomblin, Bruce. (2010b). Sequential Expectations: The Role of Prediction-Based Learning in Language. *Topics in Cognitive Science* 2, 138-153.

Perruchet, Pierre and Vintner, Annie. (2002). The self-organizing consciousness: A framework for implicit learning. In R. French & A. Cleeremans (Eds.), *Implicit learning and consciousness* (pp. 41-67). Hove, UK: Psychology Press.

Reber, Arthur, Walkenfeld, Faye, and Hernstadt, Ruth. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17, 888-896.

Robinson, Peter. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, ad Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition* 27, 235-268.

Saffran, Jenny. R., Aslin, Richard, and Newport, Elissa L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926-1928.

Sagarra, Nuria and Herschensohn, Julia. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua* 120, 2022-2039.

Sato, Mikako and Felser, Claudia. (2008). *Sensitivity to morphosyntactic violations in English as a second language.* Unpublished manuscript, University of Essex, Colchester, England.

Schmidt, Richard. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 1-32). New York: Cambridge University Press.

Schulz, Barbara. (2006). *Wh-scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of complex wh-questions* (Unpublished doctoral dissertation). University of Hawai'i, Honolulu.

Takahashi, Eri and Lidz, Jeffrey. (2008). Beyond statistical learning in syntax. In A. Gavarró & M. J. Freitas (Eds.), *Language acquisition and development: Proceedings of Generative Approaches to Language Acquisition (GALA) 2007* (pp. 446-456). Cambridge, UK: Cambridge Scholars.

Thompson, Susan P. and Newport, Elissa L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development* 3.1, 1-42.

Waters, Gloria S. and Caplan, David. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology* 49, 51-79.

Wen, Zhijun, Miyao, Mari, Takeda, Aya, Chu, Wei, and Schwartz, Bonnie D. (2010). Proficiency effects and distance effects in nonnative processing of English number agreement. In K. Franich, K. M. Iserman, & L. L. Keil (Eds.), *Proceedings of the 34th Boston University Conference on Language Development* (pp. 445-456). Somerville, MA: Cascadilla Press.

On-Soon Lee
Department of English Language and Literature
College of Liberal Arts, Korea University
145 Annam-ro, SeongBuk-gu
Seoul, 02841, Korea
Email: onsoon@gmail.com

Received: February 29, 2016
Revised version received: April 10, 2016
Accepted: April 15, 2016

APPENDIX A

C-Test

On this page you will find 2 small texts in total. Each text containing gaps where parts of some words have been left out (no whole words are missing, though). In the blanks provided, please complete words so that the sentences and texts make sense. Note that in each blank, you should only complete one word; do not add extra words. Please compete this <u>within 10 minutes</u>.

Text 1

We all live with other people's expectations of us. These are a refle_______ of th_____ trying to under____ us; th_____ are predic_____ of wh____ they th___ we will think, d____ and feel. Gene_____ we acc__ the sta____ quo, but these expec____ can be ha__ to han____ when they co____ from our fami____ and can be diff____ to ign_____ , especially wh______ they come from our par____.

Text 2

The decision to remove soft drinks from elementary and junior high school vending machines is a step in the right direction to helping children make better choices when it comes to what they eat and drink. Childhood obe____________ has bec_____ a ser_____problem in th_____ country a_____ children cons______more sugar-based fo_____ and sp_______ less ti____ getting the nece_______ exercise. Many par______ have quest________ schools' deci_____ to al_____ vending machines which disp____ candy and so_____ drinks. Many schools, tho_____ , have co______ to re______ on the mo_______ these machines generate through agreements with the companies which makes soft drinks and junk food.

APPENDIX B

Language Background Questionnaire

The questions below are intended to help us learn about your language learning experience. Your personal information will be kept confidential, and all other information will be used for research purposes only. Please read and answer all the following questions carefully. Use the blank space beside each question to clarify answers.

Name: ___________________(Gender: Female/Male)
Age: ______________ Major: ____________

1. At what age did you begin to study English? ________________
2. How many years of school instruction of English did you receive?
   (Please specify the total length _____________ year(s)
3. How many years of English grammar have you learned?
   (Please specify the total length _____________ year(s)
4. How long have you lived in a place or places where English was/is the first language of communication?
   (Please specify the total length ___________ year(s)
5. Please rate your English listening, speaking, writing, and reading abilities by circling a number on the 6-point scales below:

Listening: 1-----------------2--------------3--------------4--------------5--------------6
(beginning)(lower intermediate)(intermediate)(upper intermediate)(advanced)(near native)

Speaking: 1-----------------2--------------3--------------4--------------5--------------6
(beginning)(lower intermediate)(intermediate)(upper intermediate)(advanced)(near native)

Reading: 1-----------------2--------------3--------------4--------------5--------------6
(beginning)(lower intermediate)(intermediate)(upper intermediate)(advanced)(near native)

Writing: 1-----------------2--------------3--------------4--------------5--------------6
(beginning)(lower intermediate)(intermediate)(upper intermediate)(advanced)(near native)

6. Please rate your overall English proficiency by circling a number on the 6-point scales below:

Overall: 1-----------------2--------------3--------------4--------------5--------------6
(beginning)(lower intermediate)(intermediate)(upper intermediate)(advanced)(near native)