https://doi.org/10.24303/lakdoi.2018.26.3.51

Implicit Statistical Learning Ability in L1 Sentence Processing

On-Soon Lee
(Dong-A University)

Lee, On-Soon. (2018). Implicit statistical learning ability in L1 sentence processing. *The Linguistic Association of Korea Journal, 26*(3), 51-70. This paper aims to examine the role of implicit statistical learning ability, as measured by artificial grammar learning, in L1-English speakers' sentence processing. A total of 44 L1-English adult participants completed two artificial grammar learning tasks and two self-paced English reading tasks. One artificial grammar learning task involved a nonadjacent dependency rule and the other an adjacent dependency rule. One self-paced English reading task contained object relative clauses (a nonadjacent dependency pattern), and the other, number agreement constructions (an adjacent dependency pattern). The results show a significant correlation between the statistical learning of an adjacent dependency rule and the ability to process English number agreement, but not between the statistical learning of a nonadjacent dependency rule and the processing of English object relative clauses. The paper discusses the implications of the partial correlation, and suggests that variations in implicit statistical learning ability may explain why some learners outperform others in the performance of sentence processing.

Key Words: artificial grammar learning, (implicit) statistical learning, sentence processing, relative clauses, English number agreement

1. Introduction

The ability to track frequent patterns or regularities occurring in a language and to generalize patterns or rules is considered an important cognitive process known as implicit statistical learning. Statistical learning is the process by which language learners acquire information about distributions of elements, or patterns, in the input such as frequency or conditional probability (Aslin & Newport, 2012, p. 171; Romberg and Saffran, 2010, p. 906; Saffran, 2003).

In experimental contexts, this ability has been found to be related to various areas of language acquisition: the acquisition of word segmentation (Saffran, Aslin, & Newport 1996; Saffran, Newport, & Aslin, 1996), the acquisition of syntax-like regularities (Gómez & Gerken, 1999, 2000; Gómez, 2002; Gómez & Maye, 2005), and the acquisition of phrase structures (Thompson & Newport, 2007; Takahashi & Lidz, 2008; Onnis, Waterfall, & Edelman, 2008). In such experiments, adults and/or infants are briefly exposed, without explicit instruction, to specific patterns or rules expressed in the artificial language. The findings indicate that learners can abstract statistical patterns from a novel language. However, few studies have tested the relationship between statistical learning ability and skills involved in language acquisition. Among these few, recent research with children with Specific Language Impairment (SLI) has shown relationships between implicit statistical learning ability and language skills; for example, poor statistical learning ability is related to limited vocabulary knowledge (Evans, Saffran, & Robe-Torres, 2009) and low scores on the Clinical Evaluation of Language Fundamentals (CELF; Yim & Windsor, 2010; Yim & Rudoy, 2013). Furthermore, very recent work has found a close relationship between degree of bilingualism and implicit statistical learning ability (e.g., Onnis, Chun, & Lou-Magnuson, 2018).

Even less research has examined the role of implicit statistical learning ability in online sentence processing. While findings from a couple of studies have shown that auditory statistical learning ability is related to readers' performance in online sentence processing. Further, some methodological problems diminish their generalizability, suggesting the need for improvement in the design of statistical learning tasks. This study, therefore, attempts to examine the role of statistical learning ability in learners' performance of sentence processing, and, in addition, to assess the possibility of generalizing findings from previous studies (e.g., Farmer, Misyak, & Christiansen, 2012; Misyak & Christiansen, 2012) regardless of a task type, by employing a visual linguistic statistical learning task and two types of processing tasks.

2. Background

2.1. Relationship between Statistical Learning and Learning Rules

In an earlier study, Saffran, Aslin, and Newport (1996) investigated whether 8-month-old infants could discover the "words" in a stream of speech by manipulating the transitional probability that certain syllables would occur in specific orders in an artificial language. In the example in Figure 1, the transitional probability that one syllable follows another within a word (e.g., *pabi* or *biku*) is 1.0, whereas the transitional probability of syllable pairs (e.g., *ku go*) at word boundaries is 0.33.[1]

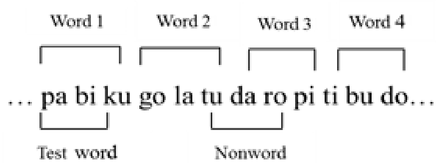

Figure 1. Stimuli used in Saffran, Aslin, and Newport's (1996) study

As the study predicted, the infants easily distinguished the "words" (e.g., *pabiku*) from the "nonwords" (e.g., *kugola*) in a test session, indicating that the infants had extracted the statistical properties of relations between syllables during the training session. The study's results suggest that this learning mechanism (i.e., statistical learning ability) may play a role not only in discerning word boundaries but also in learning grammatical relations between words. That is, if people can learn such transitional patterns from such input, then perhaps other linguistic properties (i.e., grammatical relations between words) in actual human language are learnable in this way as well.

To test this idea, Gómez and Gerken (1999) investigated whether 12-month-old infants could recognize a grammatical rule constraining the order of words in an artificial language. During the training session, the infants were briefly exposed to auditory strings of syllables (e.g., VOT PEL PEL PEL JIC) that followed the rule. The following test session evaluated their ability to

---

1) The transitional probability of Y|X =frequency of XY/frequency of X (see Saffran, Newport, and Aslin, 1996 for more information)

distinguish grammatical strings (i.e., strings that followed the rule) from ungrammatical strings. The study's results showed that the infants successfully distinguished grammatical from ungrammatical strings. This finding generalized the findings of Saffran and her colleagues by showing that infants not only were able to recognize word boundaries but also learn grammatical relations between words. A series of similar studies confirmed the results as well (Gómez & Gerken, 2000; Gómez, 2002; Gómez & Maye, 2005; Santelmann & Jusczyk, 1998).

In a more recent study, Thompson and Newport (2007) investigated whether learners can use statistical patterns to discover phrasal boundaries in an artificial language (see Figure 2). The study confirmed that statistical cues within phrase structures (e.g., AB or CD in Figure [2]) can help learners learn a phrase structure necessary to reach a certain level of syntax. Thirty-two monolingual English speakers first listened to possible sentences according to the rule in Figure 2 during a training session. The following test session showed that they were capable of distinguishing the grammatical phrasal pattern from ungrammatical phrasal patterns. The authors suggested that the ability to extract statistical patterns from an artificial language may be related to the ability to learn grammatical relations between words in the acquisition of syntax.

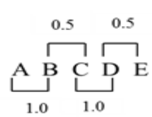

Figure 2. A phrasal rule used in Thomson and Newport's (2007) study[2)]

To extend Thompson and Newport's (2007) research, Takahashi and Lidz (2008) tested 44 native English speakers' ability to learn linear phrasal units and hierarchical phrasal structures. The results showed that the learners were sensitive enough to statistical patterns to understand both linear patterns and hierarchical structures within phrases. Taking these studies' results together, one possible conclusion is that statistical patterns within and across phrases are very

2) As Figure 2 shows, the word A always is followed by the word B within a phrase, but the word C does not always follow the word B across the phrase. Based on this, the transitional probability of two words (AB) at phrasal boundaries is 1.0, whereas the transitional probability of two words (BC) at phrasal boundaries is 0.5.

useful for learning syntactic rules.

2.2. Statistical Learning Ability and L1 Sentence Processing

A similar series of studies on statistical learning have focused on adjacent and nonadjacent dependencies between words. The main reason for focusing on these two dependency types in artificial grammar learning is that they are similar to local and long distance dependencies in natural languages. To learn a natural language's syntax, the ability to predict upcoming words or to reactivate preceding words for comprehension or production is very useful. Much work has focused on how language learners, both adults and infants, learn adjacent dependencies in structures (word segmentation in Saffran, Aslin, & Newport, 1996; syntax in Thompson & Newport, 2007 and Takahashi & Lidz, 2008). In the case of an adjacent dependency, a specific word, according to statistical probabilities, strongly predicts another, forming an adjacent (i.e., predictive) relation. For instance, a determiner (*a* or *the*) in English predicts the syntactic category of a following word as a noun.

Other researchers, on the other hand, have investigated how nonadjacent dependencies are learned. A nonadjacent dependency is a dependency between two words with at least one word intervening (Gómez & Gerken, 1999, 2000; Gómez, 2002). In other words, a specific word predicts another word that follows, but not immediately, forming a nonadjacent relation. For example, in an English object relative clause such as *the novelist that the poet admires* ___, the head noun (i.e., *novelist*) predicts its original position (marked by the underscore). Taken together, these findings from several studies imply that (i) statistical learning ability may affect learners' performance in their acquisition of natural language, and (ii) the ability to generalize statistical properties is associated with the processing mechanism of natural language; that is, the same learning mechanism applies to both artificial grammar learning and natural language acquisition. However, these conclusions remain problematic. Even though recent work has successfully tested learners' ability to abstract adjacent and nonadjacent dependencies from input, the theoretical assumption of a correlation between statistical learning ability and sentence processing has rarely been tested.

One study that has empirically tested this assumption was conducted by Misyak and Christiansen (2012), whose participants were monolingual native English speakers. They employed two artificial grammar learning tasks - one involving an adjacent dependency and the other a nonadjacent dependency - and two language comprehension tasks. The language comprehension tasks tested the participants' performance on English relative clauses and ambiguities involving phonological typicality. The study controlled for other variables involved in language learning by also assessing lexical knowledge, reading experience, verbal working memory (by using a reading-span task; Waters & Caplan, 1996), short-term memory span, fluid intelligence, and cognitive motivation. The results showed that performance in both adjacent and nonadjacent statistical learning was a predictor of language comprehension. For example, participants with higher adjacent statistical learning scores showed better accuracy in ambiguities involving phonological typicality. Participants with higher nonadjacent statistical learning scores showed better accuracy rates in their comprehension of English relative clauses. These results confirmed other studies' findings and supported the assumption that statistical learning ability and real language sentence processing ability[3] might share the same learning mechanism.

However, Misyak and Christiansen's (2012) study has some methodological problems. First, the English materials combined subject and object relative clauses, but the nonadjacent dependency rule in the artificial grammar was similar only to the long-distance dependency in object relative clauses. Previous studies in various fields have shown that object relative clauses have a deeper structural and a longer linear distance between the head noun and its gap than do subject relative clauses (considered a local dependency) (e.g., Gibson, 1998, 2000; O'Grady, 1997). Hence, while the study showed significant relations between comprehension accuracy on English relative clauses and scores on the nonadjacent dependency learning task, it would have been more useful if it had reported separate scores on the two types of relative clauses. The current study addresses this problem by using only object relative clauses. Second, the adjacent dependency rule in the other artificial grammar learning task was not

3) While the scope of processing ability varies among researchers (e.g., it can encompass general reading ability or only sentence comprehension), in this study, processing ability is used in a narrow sense to refer only to real-time sentence processing tasks.

similar to the patterns in the phonological typicality task. In the phonological typicality task, the sentences have an ambiguous word, a homonym, whose meaning is clarified by an immediately following word (e.g., bird perches$_{homonym}$ seem$_{verb}$ vs. bird perches$_{homonym}$ comfortably$_{adverb}$). However, the artificial grammar learning task used a different type of adjacent relation in artificial language; a specific word (e.g. *jux*) predicts and occurs prior to another (e.g., *dupp*), forming an adjacent relation between words (e.g., *jux dupp*). The current study attempts to address this issue by employing English number agreement between a demonstrative (*this/these*) and a noun (*house/houses*), which is more similar to the adjacent pattern presented in the artificial grammar learning task. Finally, to create similar situations for the different task types, the current study uses visual stimuli rather than auditory stimuli, as Fiser and Aslin (2002) and Yim and Windsor (2010) proved the effect of visual statistical learning including shapes on language aptitude. By using a different mode of presenting the stimuli, this study hopes to provide more convincing evidence of the role of implicit statistical learning in sentence processing. With these modifications of materials, the current study aims to examine the effect of implicit statistical learning ability on sentence processing ability. Specifically, this study posits the following research questions:

1) As a preliminary step, can the participants learn both adjacent and nonadjacent dependency rules from sequential input?
2) Does each of the two statistical learning abilities, as measured by scores on the two artificial grammar learning tasks, predict the processing ability (i.e., comprehension and grammaticality judgment) relevant to each of the two self-paced reading tasks?

3. Experiment

3.1. Participants

A total of 44 English native speakers (30 women and 14 men: mean age = 21.5, *SD* = 1.2, range 18-24), all undergraduates at an American university, participated in this experiment. Each of them completed two individual sessions on separate days.

3.2. Materials

3.2.1. Two artificial grammar learning tasks

Two artificial grammar learning tasks, one involving a nonadjacent dependency rule and the other an adjacent dependency rule, were conducted. An artificial grammar learning task typically consists of a training session and a judgment test as in Figure 3.

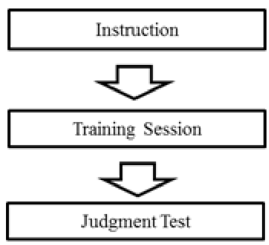

Figure 3. Procedures of an artificial grammar learning task

Participants were seated in front of a computer and informed that they would read possible sentences according to rules in an artificial language. They first completed a training session, in which they were asked to read four or five words arranged according to the rules in the artificial language (see the examples in [1] and [3]). On the computer screen, a fixation point first appeared. After it disappeared, the sentences were presented automatically, one at a time, in a random order. In the following test session, each sentence was presented in a word-by-word non-cumulative moving window paradigm, as is typical in a self-paced reading task (Just, Carpenter, & Woolley, 1982). Participants were asked to judge the grammaticality of 40 sentences (20 grammatical and 20 ungrammatical) based on what they had learned during the training session. Each of the two artificial grammar learning tasks lasted approximately 10 minutes.

For the nonadjacent dependency learning task, this study adopted the grammar from Gómez's (2002) study. The examples in (1) show the nonadjacent dependency formed between the first position (*pel, dak, vot*) and the final position (*jic, rud, tood*), with twenty-four pairs of intervening nonwords (e.g., *rob gik*)[4]. All 72 sentences generated from this grammar were shown to participants in two blocks during the training session. For the following judgment test session, ungrammatical sentences were produced by using incorrect elements in the nonadjacent dependency, as in the example (2).

(1) a. **pel** rob gik **jic**
b. **dak** vot juf **rud**
c. **vot** vev fuf **tood**

(2) a. * **pel** deg dap **tood**
b. * **dak** bov bul **jic**
c. * **vot** bup cag **rud**

For the adjacent dependency learning task, the artificial grammar was adopted from Friederici, Steinhauer, and Pfeifer's (2002) study. The examples in (3) show the adjacent dependency rule formed between adjacent words (e.g., *jux dupp*) within phrases; for example, within an A phrase (e.g., *jux dupp*), a D (e.g., *jux*) constituent always predicts and occurs prior to an A (e.g., *dupp*). In a B phrase (e.g., *tam jux dupp*), an A phrase (e.g., *jux dupp*) follows every B (e.g., *tam*). These were instantiated by nine nonsense words belonging to three lexical categories: four A members, four B members, and one D member. Thus, a sentence consists of two types of phrases, AP and BP, as in (3). All 64 sentences were generated according to this adjacent dependency rule, and then these sentences were randomly shown to participants in two blocks during the training session. During the judgment test session, ungrammatical sentences were produced by using incorrect elements in the adjacent dependency rule as in (4).

4) For training materials, 72 sentences were generated by inserting 24 pairs of intervening nonwords into each of three nonadjacent dependency pairs (i.e., *pel-jic, dak-rud, vot-tood*). Participants read these sentences twice across two blocks.

(3) a. **jux** dupp tam **jux** dupp
b. **jux** hep sig **jux** hep
(4) * **jux** dupp tam **hep** dupp

For the two types of artificial grammar learning task, during the test session, the participants read 40 sentences from the artificial language to which they had been exposed during the training session: half of the 40 sentences were grammatical, and the other half were ungrammatical. Half of the grammatical sentences were identical to sentences in the training session (i.e., 10 familiar sentences) in order to test to what extent learners had either memorized the sentences or learned the rules of the artificial language. The other half of the grammatical sentences were novel, as were all 20 ungrammatical sentences. This division into identical (familiar) versus structurally similar but not identical sentences (novel) was important to test whether the learners were capable of abstracting linguistic information beyond specific pairs of elements in order to generalize the rule.

3.2.2. Two self-paced reading tasks

The two self-paced English reading tasks were presented to the participants on a computer running E-prime (Version 2.0). Each sentence was presented one word at a time on the computer screen, left to right, in a noncumulative, moving-window manner, as the participant pushed the space bar (Just et al., 1982). Participants were asked to read as naturally as possible and then answer a yes/no comprehension question or a grammaticality judgment question. Each of the self-paced reading tasks took less than 30 minutes.

These tasks employed either sentences with object relative clauses (5a), along with a comprehension question (5b), or sentences with the adjacent number agreement in NPs (6).

(5) a. The novelist that the poet admired ____ wrote two masterpieces last year.
b. Did the poet admire the novelist? [comprehension question]
(6) a. John sold this house to his cousin last year.
b. *John sold this houses to his cousin last year.

c. John sold these houses to his cousin last year.
d. *John sold these house to his cousin last year.

Example (5a) shows a nonadjacent dependency formed between a head noun (i.e., *novelist*) and its gap (i.e., the underscore). This dependency is the same pattern as that in the nonadjacent dependency learning task (1). On the other hand, the examples in (6) show an adjacent dependency between a demonstrative (*this/these*) and a following noun (*house/houses*). For example, the demonstrative *these* predicts and occurs prior to a plural noun like *houses*; this relation is similar to that in the adjacent dependency learning task (3). In the relative clause task, which consisted of eight object relative clause sentences as in (5), with eight subject relative clauses, and 32 fillers,[5] the participants answered a yes/no comprehension question (e.g., [5b]) after reading each sentence. In the number agreement task, which consisted of 24 experimental sentences as in (6) and 48 fillers, they were asked to judge whether each sentence was grammatical or not. The ratio of yes/no answers was counterbalanced in both tasks.

3.3. Procedure and Data Analysis

All participants completed two separate sessions at least three days apart. In one session they completed the nonadjacent dependency learning task and the self-paced reading task involving object relative clauses. In the other session, they completed the adjacent dependency learning task and the self-paced reading task involving number agreement in NPs. The order of the two sessions was random. In each session, the artificial grammar learning task was followed by the self-paced reading task. Each session took approximately 40 minutes.

For the data analysis, this study follows Misyak and Christiansen (2012) in analyzing the relation between the two statistical learning abilities and the performance on the two types of syntactic structures[6]. The participants' accuracy

5) The experimental stimuli consisted of sets of subject and object relative clauses, but this study's analysis uses only the data from the object relative clauses.
6) While Misyak and Christiansen (2012) analyzed the effect of statistical learning ability on sentence processing, they looked at the relationship between judgment accuracy rates on statistical learning task and comprehension accuracy rates on processing relative clauses, because participants' judgments involve metacognitive processes.

rate on the grammaticality judgment test was calculated for each artificial grammar learning task in order to measure the participants' statistical learning ability. For the self-paced reading tasks, the participants' comprehension accuracy on comprehension questions after reading object relative clause sentences, and their accuracy on grammaticality judgments after reading number agreement sentences were measured.

3.4. Results

Table 1 provides descriptive statistics for the scores on all of the tasks. For the two artificial grammar learning tasks, the average score on the nonadjacent dependency learning task was 63.36% (*SD* = 24.27%) and that on the adjacent dependency learning task was 74.41% (*SD* = 22.14%). In both cases, the scores are significantly higher than chance level [$t(43) = 17.319$, $p < .001$ for the nonadjacent dependency learning task, $t(43) = 22.298$, $p < .001$ for the adjacent dependency learning task]. For the two self-paced reading tasks, the mean comprehension accuracy rate on the English object relative clauses was 72.7% (*SD* = 15.49%) with a mean rate of 90% for fillers. The mean grammaticality judgment accuracy rate on English number agreement in NPs was 89.2% (*SD* = 16.66%) with a mean rate of 88% for fillers.

Table 1. Descriptive statistics for the four tasks

Task	Dependent measures	Mean (*SD*)	Range (%)
Artificial grammar learning task			
a. Nonadjacent dependency	Number correct (40)	63.36 (24.27)	38-100
b. Adjacent dependency	Number correct (40)	74.41(22.14)	13-100
Self-paced reading task			
a. Object relative clause	Number correct (8)	72.7 (15.49)	38-100
b. Number agreement	Number correct (24)	89.2 (16.66)	79-100

The participants' scores on the two artificial grammar learning tasks in this study, shown in Table 1, are similar to the scores reported in Misyak and Christiansen's (2012) study for their nonadjacent dependency learning task (69.2%, *SD* = 24.7%) and adjacent dependency learning task (62.1%, *SD* = 14.3%).

The average scores from the two artificial grammar learning tasks in the current study are higher than those reported in previous studies using auditory stimuli. In this study, all the participants were sensitive enough to extract both types of rules from the artificial language. These results confirm the validity of visual statistical learning (see Fiser & Aslin, 2002; Yim & Windsor, 2010 for a review).

Table 2 shows the correlations between participants' scores on the two statistical learning tasks - one involving a nonadjacent dependency and the other an adjacent dependency - and on the two English self-paced reading tasks - one involving object relative clauses and the other English number agreement in NPs. First, scores on the adjacent dependency learning task (i.e., adjacent rule statistical learning ability) were positively related to grammaticality accuracy rates on English number agreement in NPs (r = .336). On the other hand, scores on the nonadjacent dependency learning task (i.e., nonadjacent rule statistical learning ability) were not correlated with comprehension accuracy rates on English object relative clauses (r = .184). These results are consistent with the study's predictions in the case of the adjacent dependency rule, but not in the case of the nonadjacent dependency rule. Second, scores on the nonadjacent statistical learning task are not correlated with scores on the adjacent statistical learning task. This result suggests that the two tasks tested distinct statistical learning abilities, which reflect two different learning mechanisms.

Table 2. Correlations among the four tasks

	Nonadjacent AGL task	Adjacent AGL task	ORC task	Agreement task
Nonadjacent AGL task	1			
Adjacent AGL task	-.141	1		
ORC task	.184	-.18	1	
Agreement task	.142	.336*	.243	1

Note. Nonadjacent AGL task = nonadjacent dependency artificial grammar learning task; Adjacent AGL task = adjacent dependency artificial grammar learning task; ORC task = object relative clause task; Agreement task = number agreement task.
*p < .05

Again, the main purpose of this study is to examine the relationship between the two statistical learning abilities and processing abilities on English object relative clauses and number agreement. The main question was answered by a linear regression analysis on the one significant relationship found, which was the correlation between adjacent statistical learning ability and accuracy in processing English number agreement. This analysis indicated that the scores on the adjacent dependency learning task (i.e., statistical learning ability) accounted for 9.2% of the variance in grammaticality judgment accuracy in the processing of English number agreement (F (1,42) = 5.362, p = .026, R-squared = .113, adjusted R-squared = .092). The standardized coefficient of beta showed positive beta weights, indicating that an increase in adjacent statistical learning ability led to better grammaticality judgment rates on English number agreement in NPs (β = .336, t = 2.316, p = .026). This study's finding of a partial correlation between statistical learning ability and processing ability is sufficient to suggest that the two abilities are related to each other. Furthermore, this finding resonates with the underlying assumption that statistical learning ability plays a role in language acquisition, particularly in sentence processing (see Rebuschat & Williams, 2012 for a discussion).

4. General Discussion

This study investigated whether the participants were capable of extracting the rules from each of two artificial grammar learning tasks, and furthermore examined whether an implicit statistical learning ability, as measured by the artificial grammar learning tasks, is a predictor of performance in processing sentences in natural language. For the first research question, with regard to the statistical learning of an adjacency rule, the results of this study showed that participants performed within a range of 13% to 100%. This range is consistent with the findings of some other studies; for instance, Saffran, Johnson, Aslin, and Newport (1999) found a range of 33% to 97%, and Misyak and Christiansen (2012) found a range of 40% to 97.5%. The current study also showed variations in learners' performance. With regard to nonadjacent rule statistical learning, participants performed within a range of 38% to 100%, similar to the range of 32.5% to 100% reported in Misyak and Christiansen's (2012) study. In addition, the current study showed that one-third of the participants demonstrated greater than 93% accuracy; Gómez (2002) and Misyak and Christiansen (2012) also reported that more than one-third of their participants achieved 90% accuracy. Even though the current study employed a visual artificial grammar learning task rather than an auditory learning task as used in previous studies, it showed that learners could track rules and generalize them. The findings suggest that learners are capable of extracting and generalizing rules by using statistical patterns occurring in a novel language. This study is among the first to use visual linguistic stimuli (i.e., written language, not shapes) in a statistical learning experiment. Further studies should test this method across different populations, for example, second language (L2) learners.

The second research question examined whether statistical learning ability predicts learners' performance in comprehending sentences. The study found a correlation between scores on the artificial grammar adjacency rule task and the accuracy rate on processing English number agreement in NPs. However, it found no correlation between scores on the artificial grammar nonadjacency rule task and the accuracy rate on processing English object relative clauses. These findings suggest that statistical learning ability of a specific pattern might be associated with the ability to process a particular structure in online sentence processing. Furthermore, the correlation between statistical learning and processing abilities suggest that individual differences in detecting adjacent dependencies may map onto variations in corresponding skills relevant to similar kinds of dependencies occurring in natural language. Thus, comprehending local English number agreement might rely on sensitivity to the predictive relation between a demonstrative and a noun. Previous studies have not found significant correlations between statistical learning and language skills involved in learning a natural language (e.g., Brooks, Kempe, & Sionov, 2006; Gebauer & Mackintosh, 2007), so this study's finding of even a partial correlation is noteworthy in regard to the importance of statistical learning ability as related to sentence comprehension.[7)]

7) Task differences might account for why this study found only a partial correlation between statistical learning ability and processing ability. The two artificial grammar learning tasks were judgment tasks, but the comprehension task (testing comprehension of object relative clauses) was not. This explanation is plausible, so further research will seek to improve the design of the experiment to control for this possibility.

To summarize, this study found that learners recognized statistical patterns occurring in a novel language and generalized the rules. All of the participants demonstrated some ability to learn grammatical relations between words. A significant correlation was found between statistical learning ability and processing ability in the case of an adjacent dependency. Therefore, it is possible to conclude that implicit statistical learning ability of a specific pattern may be a good predictor of performance in processing sentences involving similar specific patterns.

However, the findings present two puzzles. First, the English native speakers' comprehension accuracy rate in processing object relative clauses (72.7%) was relatively low even though their accuracy on fillers was high (e.g., 90%), which confirms that the participants were paying attention to the experimental materials. This low comprehension accuracy on object relative clauses might lead to no correlations in the case of nonadjacent dependencies, suggesting a possible avenue for further research.

Second, the study found no correlation between statistical learning and processing ability in the case of the nonadjacent dependency rule. A possible explanation is that the nonadjacent dependency in the artificial grammar learning task has no meaning, whereas comprehending object relative constructions includes interpreting form-meaning mappings. There are no such mappings in the case of an artificial grammar learning task. Another possible explanation is that the long-distance dependency in object relative clauses involves a filler and its gap - not two overt elements.[8] On the other hand, the nonadjacent dependency in the artificial grammar learning task was formed between two overt element in a pair (e.g., *pel~jic)*. In this regard, the nonadjacent pattern in the task is arguably more like the natural language dependency of a resumptive pronoun in a relative clause (e.g., the equivalent of *the man who John saw him*, which is acceptable in some language such as Hebrew). Furthermore, another factor (e.g., participants' working memory) might

---

8) The overt dependency formed between the head noun (*the novelist*) and the main verb (*wrote*) in object relative constructions (e.g., *the novelist that the poet admired wrote…*) might affect the comprehension of sentences. To control this possibility, the comprehension questions (e.g., *Did the poet admire the novelist?*) were designed to ascertain how the participants understood the clause.

lead to the partial correlation between statistical learning ability and processing ability, so this also should be examined in the future research. In spite of these issues, the findings reported here are not spurious. To the contrary, the finding of a partial correlation implies that the greater the similarity between the processing demands of artificial grammar learning tasks and the processing demands of natural language, the more likely we are to find further correlations. Thus, the study suggests that continuing to modify and improve the design of statistical learning tasks in future research will contribute to the robustness and generalizability of the findings of such research.

5. Concluding Remarks

This paper reported on the ability of adult English native speakers to abstract specific patterns (i.e., nonadjacent and adjacent dependency rules) in sequences of input in an artificial language. It then showed whether implicit statistical learning ability, as measured by the artificial grammar learning tasks, was related to performance in natural language sentence processing. The study found a close correlation between implicit statistical learning ability and the online ability to process sentences in the case of an adjacent dependency rule. However, further studies are necessary to investigate more thoroughly the suggestion made in this research across different populations and different syntactic constructions.

References

Aslin, R. N., & Newport, E. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science, 21*(3), 170-176.

Brooks, P. J., Kempe, V., & Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology. *Applied Psycholinguistics, 27*, 185-209.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 52*(2), 321-335.

Farmer, T. A., Misyak, J. B., & Christiansen, M. 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.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 458-467.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of National Academy of Sciences, 99*, 529-534.

Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit explicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 34-54

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

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. P. Marantz, & W. O'Neil (Eds.), *Image, Language, Brain* (pp. 95-126). Cambridge, MA: MIT Press.

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

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

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

Gómez, R.L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy, 7*, 183-206.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228-238.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62*, 302-331.

Misyak, J. B., Christiansen, M. H., & Tomblin, B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1:31*, 1-11.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I: Statistical learning of nonadjacent dependencies. *Cognitive Psychology, 48*, 127-162.

O'Grady, W. (1997). *Syntactic development.* Chicago: University of Chicago Press.

Onnis, L., Chun, W. E., & Lou-Magnuson, M. (2018). Improved statistical learning abilities in adult bilinguals. *Bilingualism: Language and Cognition, 21*(2), 427-433.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 906-914.

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, *44*, 493-515.

Saffran, J. R. (2003). Statistical language learning: Mechanism and constraints. *Current Directions in Psychological Science*, *12*, 110-114.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*, 27-52.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition, 69*, 105-134.

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

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development, 3*(1), 1-42.

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

Yim, D., & Windsor, J. (2010). The roles of nonlinguistic statistical learning and memory in language skills. *Korean Journal of Communication Disorder, 15*(3), 381-396.

Yim, D., & Rudoy, J. (2013). Implicit statistical learning and language skills in bilingual children. *Journal of Speech, Language, and Hearing Research, 56*(1), 310-322.

On-Soon Lee
Assistant Professor
College of General Education
Dong-A University
37 550beon-gil, Nakdongdae-ro, Saha-gu, Busan 49315, South Korea
Phone: 051-200-0924
Email: oslee@dau.ac.kr

Received on August 10, 2018
Revised version received on September 21, 2018
Accepted on September 30, 2018