INDIVIDUAL DIFFERENCES IN FIRST AND SECOND LANGUAGE SENTENCE PROCESSING: EVIDENCE FROM STATISTICAL LEARNING

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

MAY 2014

By

On-Soon Lee

Dissertation Committee:

William O’Grady, Co-chair
Luca Onnis, Co-chair
Kamil Deen
Bonnie D. Schwartz
Robert Bley-Vroman
Min-Sun Kim

Keywords: Individual differences, L2 sentence comprehension, Statistical learning, Working Memory
ACKNOWLEDGEMENTS

My sincere gratitude goes to the many people who helped me to complete this long journey. Words alone cannot express my thankfulness.

First, my deepest thanks must go to my academic advisor and dissertation committee chair, Dr. William O’Grady, who has my eternal admiration as a scholar and a teacher. I deeply appreciate his advice and his teaching, his patience and his assistance. He had faith in me, and always stood by me on my academic journey.

Next, my very sincere thanks go to my dissertation committee. Dr. Luca Onnis, as a co-chair, taught me much about statistical learning. Without his extensive knowledge of research design and his generosity in sharing his knowledge, this dissertation would not have existed. I am deeply indebted to Dr. Kamil Deen for his incalculable support and encouragement on my academic journey. With his expertise and research experience in language acquisition, he trained me to critically look at the research design as well as the data. I also owe my gratitude to Dr. Bonnie Schwartz for sharing her knowledge of second language acquisition. Her courses have inspired me and trained me to think more critically and logically. Dr. Robert Bley-Vroman’s invaluable critical comments stimulated me to improve the dissertation. Dr. Min-Sun Kim, as an external committee member, in countless ways gave me support and encouragement to successfully finish. It is due to the assistance of all of my dissertation committee members that this dissertation could be completed.

I also would like to express my gratitude to my former advisers, Dr. Kyung-Ja Park and Dr. Hikyoung Lee at Korea University, and Dr. Sungki Suh and Dr. Jong-Cheol Won (Fr. Luke) at the Catholic University of Korea, for their advice and encouragement.
Beyond my dissertation committee members, I owe my gratitude to my colleagues and friends. Many thanks to Sang-Gu Kang, Hye-Yong Kwak, Junghyoe Yoon, and Hyun-Ah Park at Korea University for actively helping me to collect the data. I am also grateful to Chae-Eun Kim, Sun-Ju Kim, and Yoon-Sang Song for their support and encouragement. They have been wonderful colleagues for discussing ideas. My appreciation goes to James Grama and Joelle Megan Kirtley for helping me with many details of the dissertation, and Joyce Hildebrand for her diligent and insightful editorial assistance. Also, Dong-Min Kim and Dr. Dong-Kwan Kong, as *sun-bae*, have been supportive, and my sincere gratitude goes to both of them.

A big thank you to friends at the University of Hawai‘i: Jung-Hee Kim, Jinsook Kim, So-Young Kim, Hee Yeoun Yoon Dennison, So-Rin Huh, Manami Sato, Apay Tracy Tang, Hiroko Sato, Da-Young Kim, Min-Yong Cho, Josephine Lee, Kum Jeong Joo, Jonny Kim, Nozomi Tanaka, Mayumi Oiwa, Yuri Lee, Kyae-Sung Park, Hyun-Ah Ahn, Kitaek Kim, Aya Taketa, Mari Miyao, and Melody Ross for sharing friendship. I also would like to express my deep appreciation to Ms. Jennifer Kanda and Mrs. Nora Lum, outstanding secretaries at the Department of Linguistics, for their extraordinary assistance.

I would like to acknowledge with gratitude the support of the Language Learning Dissertation Grant and the support of the American Association of University Women Honolulu Branch Scholarship awarded to me in the 2013/2014 academic year. For additional financial support, I am grateful to the Dr. N. H. Paul Chung Graduate Scholarship, the Center for Korean Studies Graduate Scholarship, the Sogang English Alumni Association Summer Grant, and the Graduate Student Organization at UH.
Finally, I would like to express my deepest gratitude to my parents and my younger brother (It’s your turn next!). Without their unconditional love and support, I could not have completed this dissertation.
ABSTRACT

This dissertation investigates why individuals differ in their success in learning language – in particular, why some second language (L2) learners learn more quickly and more successfully than others – by examining the correlation between ability in statistical learning tasks involving a nonadjacent dependency and success in processing filler-gap constructions in English and Korean. The results show (a) that statistical learning ability influences English native speakers’ reading times in processing English object relative clauses, (b) that an effect of statistical learning ability is observed in L2 learners’ comprehension accuracy and reading times when processing English object relative clauses, and (c) that Korean native speakers’ reading times when processing topicalized object sentences in their native language are associated with statistical learning ability. These findings shed light on the role of statistical learning in first and second language processing, implying that individual variation in this area may be one factor that explains why individual differences appear in language learning ability, at least in the case of comprehending filler-gap dependencies.
TABLE OF CONTENTS

Acknowledgment ........................................................................................................ iii
Abstract .................................................................................................................. vi
Table of Contents .................................................................................................. vii
List of Tables .......................................................................................................... ix
List of Figures ......................................................................................................... x
List of Abbreviation ............................................................................................... xi

Chapter 1: Introduction ............................................................................................. 1
  1.1 General objectives of the dissertation ......................................................... 2
  1.2 The organization of the dissertation ............................................................ 5

Chapter 2: Background .............................................................................................. 7
  2.1 Statistical learning and language acquisition ............................................. 7
    2.1.1 Statistical learning in word segmentation ........................................... 8
    2.1.2 Statistical learning and syntax ............................................................ 11
  2.2 Working memory and L1 and L2 sentence processing ............................ 18
  2.3 Individual difference in statistical learning ................................................. 25
  2.4 Research questions ....................................................................................... 31

Chapter 3: Experiment 1: English Native Speakers ................................................. 32
  3.1 Method ......................................................................................................... 33
    3.1.1 Participants ......................................................................................... 33
    3.1.2 Materials ............................................................................................ 34
      3.1.2.1 Statistical learning (SL) task ......................................................... 34
      3.1.2.2 Processing English relative clauses .................................................. 38
    3.1.3 Procedure ........................................................................................... 42
    3.1.4 Data Analysis ...................................................................................... 43
  3.2 Results ......................................................................................................... 45
    3.2.1 Statistical learning .............................................................................. 46
    3.2.2 Self-paced reading task ...................................................................... 47
    3.2.3 Combined results ............................................................................... 49
  3.3 Discussion ..................................................................................................... 55

Chapter 4: Experiment 2: Korean Learners of English ............................................. 59
  4.1 Method ......................................................................................................... 60
    4.1.1 Participants ......................................................................................... 60
    4.1.2 Materials ............................................................................................ 61
    4.1.3 Procedure ........................................................................................... 62
    4.1.4 Data Analysis ...................................................................................... 62
  4.2 Results ......................................................................................................... 62
    4.2.1 Statistical learning .............................................................................. 63
    4.2.2 Self-paced reading task ...................................................................... 65
LIST OF TABLES

Table

3.1 The artificial grammar used to assess statistical learning ability of nonadjacent dependencies ..............................................................................................................................................................................36
3.2 Sample set of experimental stimuli in English long-distance dependency experiment ..............................................................................................................................................................................41
3.3 Descriptive statistics for statistical learning task .........................................................................................................................................................................................................................46
3.4 Intercorrelations between self-paced reading task (top row) and statistical learning (left column) by English native speakers ..............................................................................................................................................................................................................................................54
3.5 Regression coefficients of two measures of SL task for two dependent measures of English ORs for English native speakers ..............................................................................................................................................................................................................................................55
4.1 Profiles of the Korean learners of English ..............................................................................................................................................................................................................................................61
4.2 Descriptive statistics for statistical learning task by L2ers ..............................................................................................................................................................................................................................................63
4.3 Intercorrelations between self-paced reading task (top row) and statistical learning (left column) by L2ers ..............................................................................................................................................................................................................................................76
4.4 Regression coefficients of two measures of SL task and C-test for two dependent measures of English ORs for all L2ers ..............................................................................................................................................................................................................................................77
5.1 Sample set of experiment stimuli in Korean long-distance dependency experiment ..............................................................................................................................................................................................................................................85
5.2 Descriptive statistics for statistical learning task by Korean native speakers ..............................................................................................................................................................................................................................................87
6.1 Descriptive statistics for L1 and L2 groups ..............................................................................................................................................................................................................................................106
LIST OF FIGURES

Figure

2.1 Results of the phrase test on day 1 (left) and day 5 (right) ........................................14
2.2 Results of the sentence test on day 1 (left) and day 5 (right) ........................................14
2.3 The phrasal tree structures for Grammars 1 and 2 .........................................................15
2.4 The sequences of mouse clicks associated with a single trial for the auditory stimulus “pel wadim rud” ........................................................................................................27
2.5 Length-adjusted reading times by each region of object relatives for ‘low’ and ‘high’ grammaticality judgment task score participants .................................................................28
3.1 English native speaker group’s mean residual RTs profile of English RCs ...................49
3.2 Mean comprehension accuracy on RCs for English native speakers .....................51
3.3 English native speakers’ mean residual RTs profile by good and poor SLers in the OR condition ........................................................................................................................52
3.4 Mean residual reading times at region 6 for English native group ...........................53
3.5 Correlation of mean grammaticality accuracy on SL task with reading times at region 6 of ORs .......................................................................................................................57
4.1 Mean comprehension accuracy by advanced and intermediate L2ers .....................65
4.2 Korean L2ers’ mean residual RTs profile for English RCs ........................................68
4.3 Mean comprehension accuracy RCs for all L2ers .........................................................70
4.4 All L2ers’ mean residual RTs profile by good and poor SLers in English ORs ......71
4.5 Mean residual reading times at region 8 for all L2ers ...............................................72
4.6 Mean comprehension accuracy on English RCs for advanced L2 group ...............73
4.7 Mean residual reading times at region 8 for advanced L2 group ..............................74
4.8 Mean comprehension accuracy on English RCs for intermediate L2 group ...........75
4.9 Mean residual reading times at region 8 for intermediate L2 group .........................76
5.1 Mean residual RTs profile of the topicalized subject and object sentences ..........89
5.2 Mean comprehension accuracy of Korean sentences for Korean native speakers ....92
5.3 Korean native speakers’ residual RTs profile by good and poor SLers in topicalized object sentences ...........................................................................................................93
5.4 Mean residual RTs at region 5 for Korean native speakers ........................................94
6.1 Mean RTs of subject and object relative clauses from related studies .................104
6.2 Mean residual RTs at the embedded verb from Experiment 1 .................................105
# LIST OF ABBREVIATION

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Accusative case particle</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative case</td>
</tr>
<tr>
<td>DECL</td>
<td>Declarative sentence-type suffix</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative case particle</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative case particle</td>
</tr>
<tr>
<td>PRS</td>
<td>Present tense suffix</td>
</tr>
<tr>
<td>RC</td>
<td>Relative clause verbal suffix</td>
</tr>
<tr>
<td>TOP</td>
<td>Topic particle</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

In this dissertation, I examine how a particular cognitive ability of individual second language (L2) learners affects their performance in the processing of unambiguous, complex structures – specifically, structures involving long-distance filler-gap dependencies. I do this by investigating a cognitive factor whose effect on L2 acquisition has never been tested: namely, statistical learning (SL) ability measured through the ability to extract grammatical rules from an artificial language that is sequentially exposed. This question is unexplored within the field of L2 acquisition and processing, and with this study, I aim to contribute to this field theoretically and empirically.

For the last two decades, studies on statistical learning have contributed to first language (L1) acquisition by showing that infants acquire words and very simple syntax by detecting the distributional cues (i.e., statistical patterns) in continuous speech streams or artificial sentences through a learning mechanism. These findings have given rise to a large body of research investigating the nature of this learning mechanism and examining the effect of SL ability on real-time comprehension. Research on L2 sentence processing has investigated how L2 adult learners comprehend sentences in real time and has found that L2 learners differ in both their processing abilities and in their degree of success in learning a second language. These studies show that L2 learners differ in terms of the speed of processing and the accuracy of comprehension: that is, L2 adult learners differ widely in their ultimate attainment. However, relatively little research has examined what factors might account for individual variation among L2 learners’ processing abilities.
The crucial interest in this dissertation is to test whether SL ability can explain this individual variation; in other words, individual differences in L2 sentence processing might be related to a cognitive capacity.

1. General objectives of the dissertation

For many years, it has been believed that all L1 learners attain a similar grammar at the end state of acquisition, whereas L2 learners differ widely in their ultimate attainment. This significant difference between L1 and L2 acquisition has been explained in terms of several factors: frequency of input, access to Universal Grammar, and general cognitive capacities. It is unclear which factor best explains this difference, but the general consensus is that differences between individual learners have a greater effect on L2 acquisition than on L1 acquisition. This observation led researchers to investigate the nature of individual difference, ultimately leading to the suggestion that there may even be variation in the linguistic proficiency of monolingual adult native speakers. Following this view, in 2012, the scholarly journal *Linguistic Approaches to Bilingualism* devoted an entire issue (2 [3]) to the theme of individual differences in native language attainment. It was argued that individual differences in language acquisition result from differences in linguistic experience. For example, more educated native speakers show better comprehension than less educated native speakers. In the same year, *Language Learning*, a scholarly journal dedicated to L2 learning, published a supplemental issue (62 [2]) that dealt with individual differences in L2 learning. One of the contributed papers (Roberts, 2012) suggested unpacking and examining the effect of factors such as working memory and processing speed/efficiency on L2 online comprehension to reach a
better understanding of L2 learners’ individual differences.

Whereas individual differences have been widely studied in L2 acquisition (for a review, see Bowden, Sanz, & Stafford [2005]) and in L1 sentence processing, only recently have L2 researchers begun to consider individual variation among L2 learners as a factor that may affect L2 processing. Despite the documentation of individual differences in language comprehension, there is no consensus on why individual differences among learners appear. Of the many possible reasons for individual differences, researchers have mostly attested the following factors as having an effect on L2 acquisition: age, proficiency, working memory,\(^1\) and prior language experience (for a review, see Larsen-Freeman & Long [1991]). Many studies have thus shown that individual factors of L2 learners might cause differences in their performance.

Even though the role of working memory in L2 sentence processing has been tested, there is no general agreement that working memory significantly correlates with L2 learners’ performance in L2 processing (see Juffs & Harrington [2011] for a review). In other words, previous studies have shown that individual differences in language comprehension cannot be fully explained by working memory capacity because the effect of working memory capacity has not been systematically shown across a variety of linguistic structures. This obscures the real picture of how cognitive capacities affect L2 learners’ performance in sentence processing.

In a very early study of L1 processing, King and Just (1991) provide evidence of individual differences in comprehension in a task in which English native speakers with

\(^1\) Working memory usually refers to L2 aptitude; Aptitude is defined in the literature as a specific talent for foreign language learning (Bowden et al, 2005, p 115). A processing approach to SLA assumes that the high working memory successfully facilitates L2 learners’ processing and acquisition by storing sequential input. Thus, most of researchers have considered it a language aptitude.

\(^2\) The transitional probability of \(Y|X = \text{frequency of } XY\)/\(\text{frequency of } X\). This is sometimes called
low working memory capacity (measured by a reading span task taken from Daneman and Carpenter [1980]) showed slower reading times at the matrix verb of object relative clauses (e.g., *The reporter that the senator attacked admitted the error*), compared to those with high working memory capacity. This difference in comprehension ability among monolingual speakers raises the possibility of other such differences, opening the door to new research questions.

This dissertation aims to address the issue of individual differences among learners in sentence comprehension by examining the relationship between learning an artificial language and processing a natural language. To test this relationship, I have adapted the artificial grammar learning paradigm for practical and theoretical reasons: (i) SL ability measures both memory and the ability to process sequential linguistic input, and (ii) there have been few attempts to correlate statistical learning and L2 sentence processing in L2 acquisition. My interest in statistical learning is motivated by two reasons. First, much of the existing literature on individual differences in language processing focuses on working memory (see chapter 2 for a review). Studies on the relationship between the role of working memory capacity and L2 learners’ online comprehension have found no robust association between working memory capacity and performance in real-time comprehension. Second, most of the L2 studies on the effect of working memory capacity have examined its effect on the resolution of temporary ambiguities, not on resolving complex structures involving long-distance dependencies that pose more of a memory load. Even so, working memory capacity failed to explain the different performance among L2 learners.

To help fill the gap on the relation between cognitive capacity and L2 sentence
processing, this dissertation includes two tasks – an SL task involving nonadjacent dependencies and an online comprehension task involving long-distance filler-gap structures. These tasks were designed to investigate (i) the degree of success with which three different groups learn abstract grammatical rules during a short exposure to an artificial language, (ii) the differences between L2 processing and L1 processing of constructions involving long-distance dependencies, and (iii) the association between L1 and L2 sentence processing ability and individual differences in SL ability.

The goal of this dissertation is twofold. First, it critically addresses the findings of previous studies on the effects of working memory on L2 processing (see chapter 2 for more details) and suggests a better way to measure learners’ cognitive capacity – SL ability. Second, it offers an explanation for individual variation in L2 sentence processing, thereby contributing to various aspects of linguistics: statistical learning, L2 acquisition, and L2 processing.

1.2 The organization of the dissertation

In this chapter, I have addressed the main purposes of this study, briefly introducing theoretical and empirical foundations that motivate it. I now provide an overview of the chapters that follow.

In the first part of chapter 2, I review research findings concerning statistical learning and language acquisition, assuming that the parsing mechanism used for statistical learning is relevant to comprehending natural language. The focus here is on findings from statistical learning studies through an artificial grammar learning paradigm; this provides the empirical evidence that adults as well as infants are capable of noticing
statistical patterns for processing nonadjacent dependencies. In the second part of chapter 2, I survey studies on individual differences in L1 and L2 sentence processing, focusing on the effect of working memory capacity. Finally, I suggest SL ability as a potential factor for the explanation of individual variation among L1 and L2 learners.

Chapter 3 introduces the methodology, presenting two tasks – a statistical learning task involving nonadjacent dependencies in the artificial language and a self-paced reading task involving English relative clauses – and then provides the experimental results, first descriptively and then statistically, from native speakers of English. Chapter 4 presents results from Korean L2 learners of English. The descriptive analysis is provided first, confirming similar findings found in previous studies on L2 processing. The rest of chapter 4 presents the combined results of all of the experiments, showing the correlation between SL ability and L2 processing. Chapter 5 introduces a self-paced reading task involving native speakers of Korean with a view to investigating the effect of statistical learning ability on the processing of Korean filler-gap constructions. This is the first test of a correlation between statistical learning and sentence processing in a language other than English. Finally, taking the results of all of the experiments together, chapter 6 provides a general discussion of the research and discusses implications for future research. I conclude the dissertation with some final remarks.
CHAPTER 2
BACKGROUND

This chapter reviews some recent work on statistical learning in infants and adults. In the first part, I briefly summarize findings in statistical learning that are relevant to language acquisition and development, emphasizing sensitivity to adjacent and nonadjacent dependencies as important linguistic properties. In particular, I focus on the assumption that the ability to learn from sequential input (i.e., statistical learning ability) has a positive effect on language acquisition. In the second part, I present findings on individual differences in L1 and L2 acquisition, focusing on the role of working memory capacity in dealing with long-distance dependencies. I then provide evidence from very recent work on statistical learning ability as a predictor of natural language performance. The chapter concludes with the research questions to be explored in this study.

2.1 Statistical learning and language acquisition

Linguists in several areas, especially in the field of language acquisition, have long been curious about how and why children are faster, more talented language learners than adults. In considering this question, researchers have empirically tested two assumptions in the statistical learning approach over the last 15 years. The first assumption is that language learners rely heavily on learning mechanisms to extract information from their linguistic environment. The second assumption is that statistical learning is a useful skill for detecting small elements (e.g., phonemes, words, and phrases) in language and segmenting a continuous speech stream into words. The last two decades have witnessed
an increasing interest in the contribution of statistical learning, including distributional
cues (i.e., conditional or transitional probabilities), to language acquisition. The following
section provides research findings indicating that statistical learning ability contributes to
language acquisition for both infants and adults.

2.1.1 Statistical learning in word segmentation

In 1996, Saffran and her colleagues published two papers regarding statistical
learning in an artificial language (i.e, an invented language). First, Saffran, Aslin, and
Newport (1996) investigated whether infants are able to discern word boundaries in an
auditory string of speech without any linguistic cue that would mark word boundaries and
pauses. They were specifically interested in whether infants are able to use transitional
probability\(^2\) to identify words in a continuous stream of speech. For example, the syllable
`pre` precedes a small set of syllables, including `tty, tend, cedes` in English. In the stream of
speech addressed to babies, the probability that `pre` is followed by `tty` is relatively high.
However, the probability that `tty` is followed by `ba`, as in `pretty baby`, is extremely low.
This difference in transitional probability is a clue that `pretty` is a word and `tyba` is not.
The ability to track transitional probabilities (which is a manifestation of statistical
learning ability) may thus be useful to infant learners in discerning word boundaries.

To test this hypothesis, Saffran and her colleagues (1996) employed the
familiarization-preferences procedure (Jusczyk and Aslin, 1995) to test 8-month-old
infants. This procedure involves two phases: familiarization and testing. In the

\(^2\)The transitional probability of Y|X = frequency of XY/frequency of X. This is sometimes called
statistical, conditional, distributional probabilities (Saffran et al., 1996b).
familiarization phase, twenty-four 8-month-old infants were exposed to unsegmented auditory strings, which consisted of four three-syllable nonsense words (e.g., bidakupadotigolabidaku...) for two minutes.

\[
\begin{array}{ccc}
.33 & .33 & .33 \\
1.0 & 1.0
\end{array}
\]

(1) ...bi da ku pa do ti go la bu bi da ku...

The strings had higher transitional probabilities within words (e.g., 1.0 for \(\text{bida}\) and \(\text{daku}\) as in [1]) and lower transitional probabilities between words (e.g., 0.33 for \(\text{kupa}\) as in [1]). In the test session, the infants were presented with two types of test stimuli, referred to as “words” and “nonwords”: exactly the same syllable combinations (i.e., words) as they had heard in the familiarization phase, and new combinations using the same syllables as in familiarization but in different combinations (i.e., nonwords). It was predicted that if the infants easily extracted statistical properties of words in familiarization, they would show different durations of listening in the two types of test items. As predicted, infants showed longer listening times for the nonwords, indicating that they extracted the relevant statistical cues even in the short exposure of linguistic experience. This study suggests, then, that 8-month-old infants learn statistical properties from sequential linguistic experience.

In another study that explored the same question, Saffran, Newport and Aslin (1996) tested adults. They created an artificial language composed of six trisyllabic words (\(\text{babupu}, \text{bupada}, \text{dutaba}, \text{patubi}, \text{pidabu}, \) and \(\text{tutibu}\)). Each word was randomly concatenated into text, with the same word never occurring twice in a row. A synthesizer
was used to produce this text with a flat intonation and identical syllable durations in order to remove any prosodic cues. The orthographic stream is represented in (2):

(2) …bupadatutibabupupatubibupadababupupidabututibu…

After listening to the speech for 21 minutes in a training session, the adult participants were asked to indicate which of two strings was more like a word. A two-alternative forced-choice test was used, in which the adults heard two trisyllabic strings: one string was a “word” according to the stream heard in the familiarization phase whereas the other was a “nonword.” Results showed that participants successfully chose the word 76% of the time, which is significantly different from the chance level of 50%. In other words, the adults, like the infants, learned (extracted) the statistical properties of the made-up language and thus were able to identify word boundaries from the sequential input in the test phase.

The findings from these two studies suggest that statistical learning – learning the pattern of certain syllables occurring together more frequently than others – plays a role not only in discerning word boundaries but also in language acquisition. That is, if people can learn most statistical property, such as a transitional pattern, from sequential input, then perhaps other linguistic properties in actual human language are learnable in this way.
2.1.2 Statistical learning and syntax

Since the work of Saffran and her colleagues, researchers have extensively adapted the artificial language paradigm in order to examine the role of transitional probabilities at the level of syntax acquisition. Statistical learning in syntax acquisition has been documented with regard to several syntactic phenomena, including syntax-like regularities (Gómez & Gerken, 1999) and phrase structure (Thomson & Newport, 2007; Takahashi & Lidz, 2008; Onnis, Waterfall, & Edelman, 2008).

Gómez and Gerken’s (1999) study, like the earlier research in learning artificial grammar, initially investigated whether 12-month-old infants could learn the very simple word order in an artificial language. For this study, infants participated in two experimental sessions. During the training session, which used the head-turn preference procedure, they were exposed to a variety of auditory grammatical strings for up to 2 minutes. In the test session, they were exposed to both grammatical and ungrammatical strings. The prediction was that infants would listen longer to grammatical strings than to ungrammatical strings, indicating acceptance and rejection, respectively. All test stimuli in this second phase of the experiment were similar but not identical to strings presented during the training phase in order to test infants’ statistical learning ability rather than their memory of strings presented during the training sessions. Test stimuli consisted of 10 grammatical strings (e.g., VOT PEL PEL PEL JIC) and 10 ungrammatical strings (e.g., *JIC PEL PEL PEL VOT). Results showed that infants successfully discriminated grammatical strings from ungrammatical strings. These results support the findings of Saffran and her colleagues that infants can extract and learn a linguistic property of an artificial language from sequential linguistic input. In addition, Gómez and Gerken’s
work shows that infants are capable of abstracting a simple structure from linguistic experience.

In a more recent study, Thomson & Newport (2007) investigated whether learners can use statistical patterns to discover phrasal boundaries in an artificial language. This study confirmed that the statistical cue within a basic phrase structure – namely, the predictive relationships between classes of words – can help learners acquire the phrase structure necessary to reach a certain level of syntax. For their study, Thomson & Newport created an artificial language that was made up of word classes A, B, C, D, E, and F. Each word class had three lexical items (a total of 18 monosyllabic consonant-vowel-consonant [CVC] nonsense words). The words were grouped into three phrases: AB, CD, and EF. If all the sentences in this language were canonical sentences as in (3), the transitional probabilities both within phrases (e.g., AB) and across phrasal boundaries (e.g., BC) would be 1.0.

\[
\begin{array}{cccc}
A & B & C & D & E & F \\
\text{transitional probabilities:} & 1.0 & 1.0 & 1.0 & 0.5 & 0.5 & 0.5
\end{array}
\]

In order to make the artificial language more like a natural language, Thomson & Newport allowed the phrase CD to be dropped as in (4), they permitted the phrase AB to appear more than once as in (5), and they allowed the phrase EF to move across AB and CD, as in (6).

(4) A B E F

(5) A B C D E F A B
Based on these syntactic features, the transitional probability between phrases (e.g., BC and DE) is reduced from 1.0 to 0.5, but the transitional probability within phrases is still 1.0. They conducted four experiments by using different types of artificial language, but I will briefly report the findings just from their first experiment.

Thirty-two monolingual English speakers completed the experiment over five days. Half the participants were exposed to only one sentence type (e.g., ABCDEF), whereas the other half was exposed to four distinct types of sentences (e.g., ABCEDF, ABCD, ABEF, CDEF). Each day, they listened to sentences for twenty minutes. Tests were conducted on day 1 and day 5 by having participants choose phrases and sentences that sounded like a phrase or sentence from this language. The results showed that both the group exposed just to the ABCDEF pattern and the group exposed to all four patterns performed better on day 5 than on day 1, as shown in Figures 2.1 and 2.2. As shown in Figure 2.1, participants in the second group learned the basic (canonical) structures of this language (e.g. ABCDEF) better than the participants in the first group. Moreover, as described in Figure 2.2, participants in the second group apparently learned the artificial language in terms of phrase structure rather than just linear strings.
These results indicate that learners are able to detect the transitional probabilities between adjacent words in sequentially presented sentences, thereby helping them to learn phrasal structure. Furthermore, learners’ statistical learning may play a role in the acquisition of syntax.

To extend Thompson and Newport’s (2007) research, Takahashi and Lidz (2008) tested whether learners are sensitive enough to statistical patterns to learn not only linear phrasal units but also hierarchical phrasal structures. To test this, two grammars were manipulated such that their canonical sentences were identical (e.g., ABCDE), but their
hierarchical structures differed, as in Figure 2.3. The big difference between the two grammars was that AB constituted a phrase (AP) in Grammar 1 but not in Grammar 2.

![Figure 2.3 The phrasal tree structures for Grammars 1 and 2](image)

Forty-four native English speakers listened to 80 sentences from each language; half of the participants were randomly assigned to Grammar 1 and the other half to Grammar 2. All participants were asked to choose the sentences that sounded like possible sentences in each language. The results showed that the learners not only learned linear patterns on the basis of various statistical patterns but also understood the hierarchical structures within phrases. Taken together, one possible conclusion is that statistical pattern within and across phrases is very useful to learning syntax.

A similar series of studies focused on two artificial grammars involving adjacent and nonadjacent dependencies. The main reason for focusing on these two dependencies in artificial grammar learning is that they are similar to local and long-distance dependencies in natural language: a very useful ability for learning natural language, therefore, is the ability to predict upcoming words or to reactivate preceding words for comprehension or production. Much work has focused on how language learners, both infants and adults, learn adjacent dependencies in structures (studies on word
segmentation, auditory stimuli presented in Saffran, Aslin, and Newport [1996]; visual stimuli presented in Fiser & Aslin [2002]; studies on syntax in Thomson & Newport [2007] and Takahashi & Lidz [2008]). A good example is the work done by Saffran and her colleagues, as discussed in the previous section. In the case of an adjacent dependency, a specific syllable or word, according to statistical probabilities, strongly predicts another syllable or word, forming a predictive relation. For instance, a determiner (*a* or *the*) in English predicts a following noun. This finding from Saffran and her colleagues’ work is also confirmed in studies on learning phrasal structures (for a review, see Thomson & Newport [2007] and Takahashi & Lidz [2008]).

Other researchers have investigated how nonadjacent dependencies are learned, a nonadjacent dependency being a dependency between two words with at least one word intervening (Gómez & Gerken, 1999; Gómez, 2000; 2002; Santelmann & Jusczyk, 1998). Although some studies failed to show infants learning nonadjacent dependencies (Newport & Aslin, 2000, 2004) because the intervening words lacked variety, in other studies, infants have successfully generalized a nonadjacent dependency between one word and another (Gómez, 2002; Gómez and Maye, 2005). For example, Gómez’s (2002) results show that both adults and 18-month-olds successfully learned nonadjacent dependencies. In the initial phase of the experiment, each learner was exposed to various strings of three nonsense words from an artificial language for 18 minutes, with a nonadjacent dependency between the first and third word in each string (i.e., *aXd, bXe, and cXf*, where the words associated with the first and third variables in each string exhibited a dependency). Four different conditions were set for the middle word, with participants exposed to 2, 6, 12, or 24 different words (i.e., the *X* variable). In the second
phase, each learner was tested as to the grammaticality of strings of words. Results indicated that (i) adults and infants were able to extract the dependency rules in an artificial language, and (ii) participants exposed to 24 different words (i.e., the $X$ variable) were able to learn the dependency rules significantly better than any other conditions.

In conclusion, recent research on statistical learning suggests that language learners are very sensitive to statistical patterns within and beyond words (Gómez & Gerken, 1999; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996; Gómez & Maye, 2005; Gómez, 2000; 2002; Onnis et al., 2004). Taken together, these findings from several studies imply that (i) statistical learning may affect learners’ performance on acquiring natural language, and (ii) the ability to generalize statistical properties is associated with processing natural language: that is, the same learning mechanism applies to both artificial grammar learning and natural language acquisition (Gómez & Gerken, 2000). 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 artificial grammar learning and language acquisition has rarely been tested.

To provide empirical evidence for this hypothesis, Misyak and her colleagues conducted several experiments that involved adapting artificial grammar learning paradigms. They also tried to explain the existence of individual differences among learners, even among adult monolingual native speakers of the same language. For a better understanding of such individual differences, the following section briefly discusses findings on individual differences in L1 and L2 sentence processing, focusing on the relation between working memory and syntactic processing. This is followed by a
review of Misyak and Christiansen’s (2012) study on individual differences in statistical learning.

2.2 Working memory and L1/L2 sentence processing

Memory constraints play an important role in sentence processing – in particular, in processing syntactically complex patterns of the sort exemplified in (7) and (8).

(7) Subject relative clauses

The reporter [who _ attacked the senator] admitted the error.

```
filler   0  gap
```

(8) Direct object relative clauses

The reporter [who the senator attacked _] admitted the error.

```
filler   1   2  gap
```

Even though the words presented in (7) and (8) are identical, (8) is perceived as more complex and more difficult than (7). Gibson (1998, 2000) explains the relative complexity of a sentence such as (8) in terms of memory resource: whereas two word that introduce new discourse referents intervene between the filler and the gap in direct object relative clause, none do in the subject relative clause. Since his explanation is important to the current research, I will briefly review it here.

According to the Dependency Locality Theory (DLT) proposed by Gibson (1998, 2000), working memory is relevant to two aspects of language comprehension: (i) storage
cost and (ii) integration cost. Gibson defines storage cost in terms of the number of syntactic heads that are necessary to form a grammatical sentence. Integration cost is defined in terms of the number of intervening new discourse referents (i.e., referential elements) between an incoming word and the syntactic head with which the incoming word is integrated.

A later and more commonly cited version of the theory proposed by Warren and Gibson (2002) focuses on the effect of distance on integration cost to explain why object relative clauses make greater demands than subject relative clauses on memory resources.

According to Gibson’s DLT, the total integration cost within dependencies between the head noun *reporter* and the position marked with underscore is greater in (8) than in (7) because of the difference in linear distance, as calculated by the number of new discourse referents that intervene between the two positions. No new discourse referent at all intervenes between the head noun (the filler, *reporter*) and the corresponding gap in the subject RC in (7), so the integration cost is zero. In the case of the direct object RC in (8), the new discourse referent (*senator*) intervenes between the head noun (*reporter*) and its gap, as does transitive verb (*attacked*) itself. The integration cost for the dependency in (8) is therefore two energy units, compared to zero for the dependency in sentence (7). Thus, the integration cost in (8) is greater than that in (7), creating a greater burden on memory resources.

Gibson’s proposal received support from several subsequent studies (Warren & Gibson, 2002; Nakatani & Gibson, 2003; Chen, Gibson, & Wolf, 2005), contributing to the idea that (1) the storage and integration costs are closely associated with the distance, 3

---

3 In other words, storage cost is associated with keeping tracking of the incomplete dependency in the current structure, and integration cost is associated with connecting an input word into the current structure.
and that (2) the memory resource is an important factor to comprehend complex sentences involving long-distance dependencies.

The literature on individual differences in sentence processing has mainly considered working memory capacity as an individual factor affecting performance in L1 and L2 acquisition. Reliable effects of working memory capacity have been reported in L1 self-paced reading studies on processing unambiguous complex structures (Gibson, 1998; Gordon, Hendrick, & Johnson, 2001; King & Just, 1991; Caplan & Waters, 1999; Just and Carpenter, 1992; Traxler et al., 2005; Pearlmutter & MacDonald, 1995). The earlier studies of working memory found that learners with a high working memory capacity comprehended syntactically complex, unambiguous sentences more quickly than those with lower capacity. Many of the studies that have tested the effect of working memory capacity have done so by employing complex sentences. In an early study of individual differences in working memory capacity and reading, King and Just (1991) found (i) that participants with lower memory capacity – as measured by the reading span task developed by Daneman and Carpenter (1980) – read more slowly than participants with higher memory capacity and (ii) that participants with lower memory capacity read even more slowly at points of high demand in terms of memory cost. In particular, participants with higher working memory capacity read significantly faster at the main verb (e.g, *was* in [9b] – a point of relatively high integrating syntactic complexity) than those with lower working memory capacity.

(9) a. The guy who __ followed the first lady was a spy.

    b. The guy who the first lady followed ___ was a spy.
In a more recent study, however, Traxler and his colleagues (2012) failed to find the effect of working memory on processing syntactically complex sentences such as those in (10).

(10) a. The director that watched the movie won a prize.
    b. The director that the movie pleased won a prize.
    c. The movie that pleased the director won a prize.
    d. The movie that the director watched won a prize.

The results showed no effect of working memory on online comprehension, indicating that the animacy of the head noun (e.g., the movie in [10c] and [10d]) in object relative clauses facilitated processing more directly than higher working memory capacity did. Based on these observations, the role of working memory as an individual factor explaining differences in L1 sentence processing performance is still being questioned. Let us turn to findings from related studies in L2 sentence processing.

Although some researchers have assumed that differences in working memory capacity should predict success in L2 learning (e.g., Skehan, 2002), little reliable effect of individual differences in working memory capacity on L2 sentence processing have been shown (for a review, see Juffs & Harrington [2011] and Roberts [2012]).

---

4 This is presumably because patterns with two animate arguments are semantically reversible: the director that the actor watched makes sense regardless of whether the director watches the actor or vice versa. Interpretation in such cases therefore requires a degree of attention to syntactic cues not required in patterns such as the movie that the director watched, in which pragmatic factors allow for just one interpretation.
In earlier work, 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 tested the hypothesis that higher scores on working memory capacity would predict better online comprehension; in particular, he looked at the question of a correlation between L2 learners’ working memory scores and their mean reading times at the critical verb (a disambiguating verb, such as looked in [11]) in temporarily ambiguous sentences.

(11) After the children cleaned the house looked very neat and tidy.

However, when learners were divided into high and low working memory groups according to their median working memory scores, no significant correlation was found between individual variations in reading times at the critical region (looked) and working memory. Juffs pointed out one weakness of his study in that the different L2 groups were not homogeneous in terms of their English proficiency. In particular, the Japanese L2 group was less proficient than the other L2 groups. The different levels of English proficiency, therefore, may have masked any effects of working memory capacity on online processing.

In more recent work, Havik et al. (2009) tested for an effect of working memory capacity on the processing of Dutch relative clauses; that is, they investigated whether German speakers who were advanced learners of Dutch and had a higher working memory capacity showed a processing advantage over advanced learners with a lower
memory capacity in reading temporarily ambiguous Dutch subject relative clauses, as in (12a), compared to dispreferred object relative clauses, as in (12b).

(12) a. Subject relative clauses
Daar is de machinist die de conducteurs heeft bevrijd uit het brandende treinstel.
Literally, ‘That is the engine-driver who the guards has saved from the burning train-carriage.’
(English translation: That is the engine-driver who has saved the guards …)

b. Object relative clauses
Daar is de machinist die de conducteurs hebben bevrijd uit het brandende treinstel.
Literally, ‘That is the engine-driver who the guards have saved from the burning train-carriage.’
(English translation: That is the engine-driver who the guards have saved …)

The results showed no effect of working memory capacity on L2 learners’ on-line sentences processing, but there was an effect of working memory capacity on learners’ judgments to determine whether experimental sentences were true or false: advanced German learners of Dutch with a high working memory capacity performed like native speakers of Dutch with a low working memory capacity.

In another study, Williams (2006) found an effect of working memory capacity on the processing of English wh-questions, such as (13), in Korean, Chinese, and German L2 learners of English.

In Dutch, the temporary ambiguity between subject and direct object relative clauses occurs before encountering the auxiliary verb (heeft/hebben); only number agreement between the auxiliary verb and a preceding noun resolves the temporary ambiguity.
During the presentation of sentences via word-by-word self-paced reading, participants were asked to do a “stop-making-sense” decision. Findings indicated that L2 learners with a high working memory capacity behaved like native speakers, showing an earlier response to an implausible object for the verb (e.g. *which river – push*) than to a plausible object for the verb (e.g., *which girl – push*).

In summary, findings from studies of L2 learners are not clear enough to conclude that working memory capacity plays an important role in the speed of reading sentences by L2 learners. In addition, the experimental sentences used in L2 studies have not been designed to look directly at the relation between the effect of working memory and the processing of syntactically complex structures as they have in L1 studies. To confirm the effect of working memory for individual differences in L2 sentence processing, appropriate structures must be employed.

Moreover, the reading span test popularly employed in L2 sentence processing to examine the effect of working memory might be problematic for two reasons: (i) the reading span test fails to measure the processing cost of incoming words (Ariji et al.,

---

6 Participants were asked to respond as soon as they thought that the sentence had stopped making sense by pressing a space bar. After any stop making sense decision, the participants continued reading the remainder of sentence by pressing the mouse button.

7 Daneman and Carpenter (1980) designed a reading span test, which aimed to measure both storage and processing costs. For the procedure of a reading span test, the experimenter shows participants a set of sentences written on a sheet of paper, which they are asked to read aloud. After reading all of the sentences, they are asked to recall the final word of each sentence. The number of sentences increases in each set. The reading span size is defined as the maximum number of sentences for which participants are able to recall the final word.
2003), and (ii) this task focuses on memorizing the final words of sentences, not on reading sentences (Waters & Caplan, 1996).

Based on these observations, recent work on statistical or artificial language learning has tested the idea that if working memory fails to fully explain individual differences, statistical learning ability, which measures memory as well as processing, might be a potential predictor of language learning success. A few researchers have adapted the artificial grammar learning paradigm in an attempt to explain why adult monolingual native speakers of English show different performance in language comprehension (Wells et al., 2010; Misyak and Christiansen, 2010, 2012; Farmer et al., 2012). Their studies are based on the hypothesis that variation in artificial grammar learning ability will predict different processing ability in natural language. The great advantage of using artificial language is the ability to control the input and exclude the influence of other factors (e.g., reading experience or prior learning) that may affect the learning of natural language. The following section presents findings from studies related to individual difference in statistical learning.

2.3 Individual difference in statistical learning

Recent work on statistical learning ability in the artificial grammar learning paradigm has been designed to examine empirically the correlation between learning ability in artificial grammar learning and processing ability in real-time sentence processing (Misyak et al., 2010; Misyak and Christiansen, 2012). To test this possibility, Misyak et al. (2010), in order to clarify the potential role of statistical learning ability in processing language, adapted an artificial grammar learning paradigm that involves nonadjacent
statistical dependencies. The key assumption underlying their work is that nonadjacent statistical dependencies resemble those found in relative clauses in a limited but relevant way: they hold across intervening material. See chapter 6 for further discussion of this point.

Misyak et al.’s (2010) research was based on two major experiments. In a within-design experiment, twenty monolingual native English speakers completed two experiments in one day. The first experiment employed the same artificial grammar containing nonadjacent dependencies as was used in Gómez’s (2002) study. A made-up language had three forms (i.e., aXd, bXe, and cXf), with initial and final items forming a nonadjacent dependency pair: a with d, b with e, and c with f. Beginning and ending stimulus tokens (a, b, c; d, e, f) were represented by pel, dak, vot and rud, jic, tood; middle tokens (X) were represented by 24 disyllabic nonsense words⁸: wadim, kicey, puser, fengle, coomo, loga, gople, taspu, hiftam, deecha, vamey, skiger, benez, gensim, feenam, laeljeen, chila, roosa, plizet, balip, malsig, nolbo, and wiffle. Participants were also exposed to ungrammatical forms – *aXe, *aXf, *bXf, *cXd, and *cXf – in order to test their learning in a test phase.

All participants followed the same procedure. A computer screen was partitioned into a grid consisting of six equal-sized rectangles: the leftmost column contained the beginning items (a, b, c), the center column the middle items (X₁…X₂₄), and the rightmost column the ending items (d, e, f), as in Figure 2.4. After 250 msec. of familiarization to the six visually presented nonsense words, the auditory stimuli were

---

⁸ Results from Gómez (2002) and Onnis et al. (2004) showed that high variability of the intervening element (e.g., 24 items rather than 2 or 12) helped participants to learn the nonadjacent dependency (e.g., aXd) from sequential input.
played. Participants were instructed to use a computer mouse to click upon the rectangle with the correct nonsense word as soon as they heard it.

Figure 2.4 The sequence of mouse clicks associated with a single trial for the auditory stimulus “pel wadim rud” (from Misyak et al., 2010)

Participants were exposed to a total of 432 strings (72 unique strings × 6 training blocks): each training block involved with 72 unique strings (24 strings × 3 dependency pairs). After being exposed to these 432 strings, participants were presented with 24 ungrammatical strings. The grammaticality judgment task, which tested the participants on 12 strings (six of which were ungrammatical), followed: they were instructed to endorse or reject each string according to whether they judged it to follow the rules. Results from the first experiment showed that the participants’ test accuracy scores averaged 58.2% and reflected considerable individual variation.

In the second experiment, participants read 40 experimental English sentences (20 with subject relatives clauses [SRs] and 20 with direct object relative clauses [ORs]), as in (14), along with 48 filler sentences; a Yes/No comprehension probe followed each sentence.

(14) a. SRs: The reporter that attacked the senator admitted the error.

   b. ORs: The reporter that the senator attacked admitted the error.

---

9 This is a questionable feature of Misyak et al.’s design, as it has no counterpart in the acquisition of a language by children in a natural setting.
All sentences were presented using word-by-word window-moving self-paced reading (Just, Carpenter & Woolley, 1982). The overall accuracy rate of comprehension was 88.8%, and, consistent with the previous studies, comprehension for object relative clauses (80.5%) was poorer than that for subject relative clauses (87.2%). To test whether statistical learning ability mediated individual differences in corresponding reading time patterns in the comprehension task, participants were classified as ‘low’ or ‘high’ in statistical learning ability according to their grammaticality judgment task scores in Experiment 1 (with 50% as the cut-off level). Taken together, while the two groups (i.e., high grammaticality judgment task score participants vs. low grammaticality judgment task score participants) did not differ in their processing of subject relative clauses, reading times diverged at the main verb (i.e., admitted as in [14b]) of object relative clauses, as shown in Figure 2.5: participants in the high grammaticality judgment task score group showed less processing difficulty for reactivating a preceding noun (i.e., reporter as in [14b]) in order to keep track of a long-distance dependency in a sentence.

![Figure 2.5](image.png)

Figure 2.5 Length-adjusted reading times by each region of object relatives for ‘low’ and ‘high’ grammaticality judgment task score participants (derived from Misyak et al., 2010).
These findings suggest that statistical learning ability from sequential input is related to processing ability of complex syntactic structures.

In a later study, Misyak and Christiansen (2012) added one more artificial grammar (i.e., a grammar including an adjacent dependency) to test whether statistical learning ability is a better predictor of online sentence comprehension than verbal working memory span scores. To compare the role of statistical learning ability with that of verbal working memory span, they employed two statistical learning tasks that included adjacent and nonadjacent dependencies along with three language comprehension tasks involving (i) subject relative clauses (e.g., *The reporter that attacked the senator admitted the error*) and direct object relative clauses (e.g., *The reporter that the senator attacked admitted the error*); (ii) clauses with animate/inanimate noun constructions (e.g., *The defendant examined by the lawyer turned out to be unreliable* or *The evidence examined by the lawyer turned out to be unreliable*); and (iii) noun/verb homonyms with phonologically typical or atypical noun/verb resolutions (e.g., *Chris and Ben are glad that the bird perches [seem easy to install] /[comfortably in the cage]* or *The teacher told the principal that the student needs [were not being met] /[to be more focused]*). Two artificial grammars each were adapted from Gómez (2002) and Friederici et al. (2002).

Additionally, a verbal working memory span task was conducted, including extra measurements of language aptitude related to lexical knowledge, reading experience, short-term memory span, fluid intelligence, and cognitive motivation.

Results from Misyak and Christiansen (2012) show that performance in both adjacent and nonadjacent statistical learning was a better predictor of sentence comprehension than verbal working memory span scores were. Statistically, 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 of comprehension of relative clauses. In other words, high learning scores in adjacent and nonadjacent statistical learning predicted better performance in related language tasks – ambiguities involving phonological typicality and comprehension of relative clauses, respectively. This conclusion confirmed the previous findings by Misyak and her colleagues that supported the assumption that statistical learning ability and processing ability in real language sentence processing might share the same learning mechanism. Furthermore, these findings from monolingual native English speakers suggest that there are substantial differences in adult native speakers’ comprehension abilities.

This raises an important question: Are differences in English native speakers’ performance caused by differences in linguistic experience? This also relates to the question of why L2 learners reach different end states of their target language. Even though various language aptitudes (e.g., verbal working memory) have been studied in L2 acquisition, higher statistical learning ability could be a better predictor of language learning speed and aptitude. It is valuable, therefore, to look at the role of statistical learning ability in online language comprehension. Thus, I have extended the studies summarized above from English to Korean, including both L1 speakers and L2 learners, in order to add more empirical evidence supporting the hypothesis that statistical learning ability is a good predictor not only of different linguistic abilities in monolingual speakers but also of individual differences in L2 learners’ attainment.
2.4 Research questions

This dissertation will extend the scope of this research in terms of a different language and a different population. To test the possibility of a correlation between statistical learning and natural language acquisition, my research employs an artificial grammar learning task involving nonadjacent dependencies, which correspond to long-distance dependencies in natural language (see Gómez [2002]). The logic underlying my study is that success in learning an artificial grammar requires an aptitude for observing and learning dependencies – that is, correlations between the presence of one item and the occurrence of another. Work by Misyak et al. (2010) and Misyak and Christiansen (2012) suggests that statistical learning ability is required for learning long-distance dependencies. The key objective of my research is to correlate statistical learning ability with variation in language processing involving long-distance dependencies. In this regard, I extend the work of Misyak and her colleagues on the processing of English by native speakers to the processing of Korean by native speakers and the processing of English by Korean L2 learners. Specifically, this dissertation will explore the following research questions:

1) Does performance on statistical learning tasks correlate with accuracy on comprehension tasks?
2) Does performance on statistical learning tasks correlate with reading speed at critical regions in the processing of filler-gap dependencies in patterns such as relative clauses?
CHAPTER 3
EXPERIMENT 1: ENGLISH NATIVE SPEAKERS

This chapter presents an experiment consisting of two tasks that were completed by English native speakers, one involving statistical learning (SL) and the other involving the processing and comprehension of relative clauses. The experiment was designed to investigate the relation between statistical learning and the online comprehension of relative clauses by: (a) comparing the comprehension of object relative clauses between good and poor statistical learners (SLers), whose statistical learning ability is determined on the basis of statistical learning scores; and (b) comparing the reading times at the vicinity of the main verb in object relative clauses for good and poor SLers. After describing the methodology, I report the results from the two tasks. I begin with a descriptive analysis and then present the combined results of the two tasks, with inferential statistics. The results of these experiments were used as a baseline for testing the relation between statistical learning ability and the processing of English relative clauses by Korean learners of English (L2ers). The results of that experiment are reported in the next chapter.

This experiment, conducted with English native speakers, was designed to replicate previous studies that explored the relation between statistical learning ability and the processing of English relative clauses. The research questions addressed in this chapter are as follows:
a. Do English native speakers track/learn nonadjacent dependencies in a statistical learning task involving an artificial grammar, as reported in previous studies of English native speakers?

b. Do English native speakers show difficulty in processing object relative clauses on the basis of comprehension accuracy rates as well as reading times at the critical regions (i.e., the vicinity of main verb), as confirmed in previous studies on L1 processing?

c. Is the effect of statistical learning ability observed in processing English object relative clauses? It is predicted that higher grammaticality accuracy rates in the statistical learning task will correlate with higher comprehension accuracy rates in processing English object relative clauses and faster reading times at the critical regions.

3.1 Method

3.1.1 Participants

Fifty-three native English speakers (30 women and 23 men; mean age = 22.5, SD = 1.2, range = 18–26), all University of Hawai‘i undergraduates, participated for course credit or money. Each of them completed both tasks in one session: a statistical learning task and a self-paced reading task. In order to look at the individual differences, we did not exclude any participants from the data analysis.
3.1.2 Materials

This section describes (a) the statistical learning task, including a nonadjacent dependency; and (b) the self-paced reading task, including English relative clauses. The aim of the current study was to investigate the effect of statistical learning ability – specifically, English native speakers’ sensitivity to nonadjacent dependencies – on the processing of English object relative clauses.

3.1.2.1 Statistical learning (SL) task

We employed a statistical learning task, implementing a type of an artificial grammar, in order to measure learners’ sensitivity to a nonadjacent dependency between constituents – the so-called statistical learning ability in this study. The artificial grammar that we used has clear parallels within natural language structures. Studies have shown that statistical learning ability may be important for processing long-distance dependencies (e.g., Gómez, 2002): that is, the ability to track remote dependencies between two constituents is a fundamental linguistic ability. Recent work has also suggested that learning nonadjacent dependencies in artificial grammar is relevant for understanding natural language (Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006).

The auditory stimuli used for the statistical learning task were typical of those used in other studies to assess statistical learning ability. In the current study, however, all stimuli for the statistical learning task were presented visually in order to create a situation similar to the one encountered in the sentence processing tasks. Previous studies have shown that adults are sensitive to visual stimuli in statistical learning tasks (see
Altman et al., 1995; Tunney et al., 1999). In particular, all sentences in the statistical learning task were presented using a word-by-word paradigm, thus presenting similar circumstances to those in sentence processing tasks in which the participants took part.

Traditionally, a statistical learning task consists of a training (or familiarization) phase and a test phase. Before starting this task, participants were informed that they would read a series of sentences from a made-up language, all of which are possible sentences according to the artificial grammar. They also knew that each sentence would consist of exactly four words, arranged grammatically according to the rules of the imaginary language. Participants were told that they would then be tested in order to see how much they had learned from the first phase. This nonadjacent statistical learning task was conducted on a computer running E-prime (Version 2.0). In the training session, each sentence was presented automatically. For a 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 the space bar (Just, Carpenter, and Woolley 1982). In the training session, each sentence was presented after a fixation point appeared on the screen for 250 milliseconds. Then, each word of a particular sentence appeared automatically for 250 milliseconds (ms) on the left side of the computer screen, and then the sentence consisting of these four words was held for a second on the computer screen. Other sentences followed randomly with the same presentation method. The training phase lasted approximately 10 minutes.

After finishing the training session, participants were shown some sentences, one at a time, and asked whether they thought they were possible sentences in the made-up language. Before the actual test session, participants were given a practice session in
which they were instructed to say which of 4 sentences presented were more like possible sentences in a made-up language. During the test session that followed, they read 40 sentences from the made-up language to which they had been exposed during the training phase: 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, after the short exposure time in the training phase, 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.

Table 3.1 The artificial grammar used to assess statistical learning ability of nonadjacent dependencies.

<table>
<thead>
<tr>
<th>The artificial grammar</th>
<th>Examples of training items</th>
<th>Examples of test items</th>
</tr>
</thead>
<tbody>
<tr>
<td>S: aXYd, bXYe, cXYf</td>
<td>aXYd pel rop gik jic</td>
<td>aXYd pel rop gik jic</td>
</tr>
<tr>
<td>X: {X₁, X₂, …X₂₄}</td>
<td>bXYe vot poy juf tood</td>
<td>*aXYe *pel bup cag tood</td>
</tr>
<tr>
<td>Y: {Y₁, Y₂, …Y₂₄}</td>
<td>cXYf dak vev fuf rud</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* * indicates ungrammatical sentences in this artificial language

For the nonadjacent dependency learning task, this study adapted the grammar with minor modification from Gómez’s (2002) study. That study was based on three sets of dependency pairs (i.e., a-d, b-e, c-f); each string contained one intervening element in the middle of the dependency pair (i.e., aXd, bXe, cXf). In the current study, consistent with the nonadjacent dependency grammar in Table 3.1, an additional intervening element was inserted between the members of the dependency pairs (i.e., aXYd, bXYe, cXYf). The
three pairs created between the first position (a, b, c) and the final position (d, e, f) formed nonadjacent dependencies similar to the long-distance dependencies in English relative clauses; the dependencies of the SL task resemble those found in relative clauses in a relevant way - they hold across intervening words. These three pairs of nonadjacent dependencies were instantiated with monosyllabic nonsense words (dak, pel, vot; jic, rud, tood).

Previous work has shown that statistical learning of nonadjacent dependencies is facilitated for both infants and adults by a high variability in the intervening elements between the two elements forming nonadjacent dependencies (Gómez, 2002; Gómez & Maye, 2005; Onnis et al., 2003; Onnis et al. 2004). Thus, each of the intervening elements – X and Y – was drawn from 24 monosyllabic nonsense words (X: bov, bup, deg, doz, dyt, fet, fok, gol,gos, gug, kav, lat, lyz, pag, poy, rop, seb, siz, suv, tib, ven, eve, wat, zaf; Y: bul, cag, dap, fud, fuf, gaz, gik, gof, gyb, juf, kem, kes, kif, kuc, lec, lig, lod, lut, mup, ned, nim, nom, pez, tol). This amount of variability contributes to successful learning in SL tasks. For instance, suppose that three types of grammatical sentences contain the following sequence of category constituents: aXYd, bXYe, cXYf (e.g., dak bov bul jic, pel bup cag rud, vot deg dap tood). All 72 sentences generated from this grammar were shown to participants in two blocks during the training session; all participants were exposed to a total of 114 sentences. For the test session, ungrammatical sentences were produced by using incorrect final elements in the nonadjacent dependency, as in Table 3.1. Thus, ungrammatical sentences were formed with three set of pairs: for example, *aXYe, *bXYf, *cXYd (e.g., *dak bov bul rud, *pel bup cag tood, *vot deg dap jic).
3.1.2.2 Processing English relative clauses

In the self-paced reading task, English subject and direct object relative clauses were used. The object relative clause is a good choice for investigating whether learning nonadjacent dependency is relevant for comprehending long-distance dependencies within natural language.

I begin by reviewing the difficulty of processing object relative clauses. For several languages, the greater difficulty of direct object relative clauses (ORs) over subject relative clauses (SRs) has been examined in many areas of language acquisition and psycholinguistics, including the role of memory in language comprehension (in L1: Gibson, 1998; Gordon, Hendrick, & Johnson, 2001; King & Just, 1991; Lewis et al., 2006; in L2: Felser & Robert, 2007; Havik et al., 2009; Juffs & Harrington, 2011; Rodríguez, 2008), typical and atypical child language development (Booth, MacWhinney & Harasaki, 2000; Friedmann & Novogrodsky, 2007; Kidd, Brandt, Lieven & Tomasello, 2007), and individual differences in adults (Just and Carpenter, 1992; King and Just, 1991).

Although several explanations have been offered for the greater difficulty of ORs, no consensus has been reached among researchers. It may be that the greater distance between the “filler” (e.g., the novelist in [1]) and the “gap” (the underscores in [1]) makes ORs more difficult than SRs for reasons relating to processing and memory cost.

(1) a. SR: The novelist that ____ admired the poet wrote two masterpieces last year.
   
   b. OR: The novelist that the poet admired ____ wrote two masterpieces last year.
For example, Gibson and colleagues (Gibson, 1998, 2000; Grodner & Gibson, 2005; Warren & Gibson, 2002) have suggested that the relative difficulty of comprehending different types of sentences, especially in online comprehension, is related to two components: (1) integration cost associated with connecting an incoming word into the current structure and (2) storage cost associated with keeping track of long-distance dependencies. In this process of keeping track of the dependency between an upcoming word and the earlier word, a more distant word is thought to be more difficult to reactivate than a more recent word. According to this dependency locality theory, an SR is easier to comprehend than the OR because of the relatively greater distance in an OR between a head noun (novelist in [1]) and the gap position (underscores in [1]). The relation between a filler and its gap in an OR is called a long-distance dependency. As proposed by the dependent locality theory (Gibson, 1998) and the theory of working memory retrieval (Lewis & Vasishth, 2005), the greater number of intervening words in an OR increases processing load during the retrieval process of lexical and grammatical knowledge. Thus, a relatively longer reading time at the vicinity of the main verb, the so-called critical region, will be observed in the OR condition as compared to the SR condition.

To examine the effect of statistical learning ability on learners’ processing of ORs, I used a self-paced reading paradigm, providing two dependent measures: (a) comprehension accuracy and (b) reading times at particular regions within the sentences. Such a paradigm is based on two assumptions. First, the additional memory load needed

---

10 Gibson (1998) defines storage costs in terms of the number of syntactic heads that are predicted so as to form a complete grammatical sentence. Integration costs are defined based on the number of intervening discourse referents between the newly integrated word and the syntactic head with which the new word is integrated (Gibson, 2000).
to keep track of the dependency between an incoming word and the current words causes processing difficulty. Example (2) shows an OR in which the displaced phrase (the novelist) is extracted from the original position (underscore).

(2) The novelist that the poet admired __ wrote two masterpieces last year.

If participants recognize the potential gap position of the phrase the novelist at the critical word (i.e., wrote), they are likely to slow down at the critical region. Longer reading times reflect processing difficulties at that region. This method is thus able to test processing difficulty by recording reading times across all regions. Second, additional data is gathered involving comprehension questions. For example, participants were asked to read the sentence in (2) and then to answer a comprehension question (e.g., Did the poet admire the novelist?). The accuracy analysis from these data shows how well participants comprehended sentences.

The relative clauses processing task was conducted on a computer running E-prime (Version 2.0). Each sentence was presented one region (i.e., word) at a time on the computer screen, left to right, in a noncumulative, moving-window manner as a participant pushed the space bar (Just, Carpenter, and Woolley 1982). Participants were asked to read as naturally as possible and then answer a yes/no comprehension question to ensure that they had attended to the stimuli. All reading times for each region were recorded, along with responses to comprehension questions. Both tasks together lasted less than 30 minutes for each participant.
For the self-paced reading task involving long-distance dependencies, the experiment involved 16 sets (2 sentences each) of experimental stimuli, as illustrated in Table 3.2. Each stimulus contained 11 regions (i.e., words). The SR is termed ‘subject’ relative clause because the noun modified by the relative clause – that is, the head noun (the novelist in Table 3.2) – is the subject of the relative clause’s verb (admired). The OR is termed ‘direct object’ relative clause because the noun modified by the relative clause – that is, the head noun (the novelist) – is the direct object of the relative clause’s verb (admired). The semantic bias of embedded verbs (admired) in all sentences was manipulated, such that the two animate nouns were reversible; for example, in the sentences in Table 3.2, it is semantically feasible that the poet could admire the novelist and vice versa. Also, the animacy of head nouns was controlled in order to remove the cue for direct object relative clauses: the head nouns of direct object relative clauses are more frequently inanimate than animate (Reali & Christiansen, 2007). We call region 7 (wrote) the critical region because this is where the reader has the opportunity to look for the original position (the gap) of the filler (the novelist): the critical region is where the relative slowdown of reading time is expected to appear because of an increase in the processing load in order to keep track of a long-distance dependency.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR</td>
<td>The novelist that admired the poet wrote two masterpieces last year.</td>
</tr>
<tr>
<td>OR</td>
<td>The novelist that the poet admired wrote two masterpieces last year.</td>
</tr>
</tbody>
</table>
The regions of primary interest are 6, 7, and 8, where participants are likely to slow down to keep track of the dependency. The 16 sets of experimental stimuli were distributed in a Latin square design across two lists, randomly assigned to participants so that they each saw only one condition of each experimental item. Each list contained 16 sentences containing relative clauses (8 SRs and 8 ORs), along with 32 fillers. All fillers were complex sentences: 16 conjoined constructions (e.g., *Steven called Jennifer and invited her to be in his film*) and 16 bi-clausal constructions (e.g., *Even though Tom drove his car slowly, he got a speeding ticket*). In total, participants read 48 sentences. After reading each sentence, participants were asked to answer a yes/no comprehension question. The ratio of yes/no answers was counterbalanced.

3.1.3 Procedure

Each English native speaker completed both tasks in one session; all participants completed the statistical learning task first, followed by the self-paced reading task. Participants were asked to sit in front of a computer screen. The statistical learning task consists of two phases. In the training phase, participants were informed that they would be asked to read several sentences of a made-up language, which were automatically presented word by word through E-prime software over a 10-minute period. During the test phase, they read sentences from the made-up language one word at a time with each press of the space bar on the keyboard, and judged whether the sentences were possible or not.

For the self-paced reading task, participants were asked to read English sentences word by word as they pressed a space bar; after each sentence, they were asked to give a
yes/no answer to a comprehension question such as (3b) (see Appendix C). A language background questionnaire followed (see Appendix G). In total, each individual session lasted approximately 45 minutes.

(3) a. The professor that criticized the student went on vacation in Texas.
   b. Did professor go on vacation in Texas? Correct Answer: Yes

3.1.4 Data analysis

In a statistical learning task, participants’ statistical learning ability is calculated according to the grammaticality accuracy rates, along with reading times at the critical region (i.e., the last word of the sentence in the artificial language). Since there were 40 sentences (20 grammatical and 20 ungrammatical) in the test phase, the total maximum correct responses was 40 if participants accurately judged whether test sentences were grammatical or ungrammatical on the basis of what they learned during the training session. The percentage of correct responses on both familiar and novel sentences was also calculated so as to examine the effect of novelty on the accuracy rates. For the supplementary analysis, reading times at the last word of the SL task, where the integration cost may be high because of the need to keep track of the nonadjacent dependency between the first and last words, were recorded. It also tested the potential correlation between the reading times at the last word and the speed of reading times in the self-paced reading task. This analysis will be referred to as the baseline, replicated with findings of previous studies on statistical learning.

For the self-paced reading task, I calculated two dependent measures: (a) the accuracy of comprehension questions and (b) the reading times at critical regions (i.e.,
disambiguating regions). I focused in particular on the comprehension accuracy and reading times of object relative clauses, because in previous studies, uniform ceiling-like behavior has been observed for subject relative clauses. The comprehension accuracy rate was calculated on the basis of correct responses to comprehension questions.

All analyses of reading times included those on which the comprehension question was answered correctly. Reading times (RTs) were calculated according to the following procedure. To eliminate RT outliers, raw RTs were trimmed in the following way. First, any RT longer than 2500 ms was replaced with 2500 ms. Next, RTs more than ±2.5 standard deviations above or below the mean were removed. Any RTs longer or shorter than ±2.5 standard deviations above or below the mean were replaced with those points. Then, residual RT transformations were done on the raw RTs to factor out effects of differences in word length and reading speed. Residual RTs for each participant were obtained by calculating each region’s expected RT (from a linear regression equation according to word length) and subtracting it from the raw RTs (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). If a raw RT equals its expected RT, the residual RT is zero; if the raw RT is longer than expected, the resulting residual RT is positive, and if shorter, negative. This analysis will be referred to as the baseline, replicated with results of sentence processing.

For the comparison between statistical learning and online comprehension, the participants were divided into two groups – good statistical learners (good SLers) and poor statistical learners (poor SLers) – according to their grammaticality accuracy rates. The cut-off value was defined as 90 out of a maximum 100%, and participants whose scores were above or equal to 90% of grammaticality accuracy were categorized as good
Participants whose scores were below 90 were classified as poor statistical learners. We opted for this value because the traditional way of splitting the group by its median in previous L1 and L2 processing studies typically results in rather homogeneous groups and in misclassification at the margin, so that effects of the dividing variable are often not observed (for discussion, Hopp 2013; Conway et al., 2005).

In order to see whether participants’ statistical learning ability affected the processing of object relative clauses, we will look at difference in comprehension accuracy and reading times at the critical regions between good and poor statistical learners, by using independent t-tests and linear regression. These results from the statistical analysis will provide the empirical evidence in order to examine the hypothesis that the better learners on the statistical learning task responded more accurately and were faster in sentence processing. For all the statistical tests reported in this dissertation, the alpha level was set at .05.

3.2 Results

This section reports results from the English native speakers, all of whom completed both the statistical learning task and English relative clause task. For the statistical learning task, the analyses were conducted on the dependent measures of participants’ grammaticality accuracy and reading times (RTs) on the last word. For the self-paced reading task, analyses were conducted on the dependent measures of comprehension accuracy rates and reading times at the critical regions. We will now look the results from both tasks; this is followed by a comparison between the results of the two tasks.
3.2.1 Statistical learning

Table 3.3 shows the descriptive statistics for the statistical learning task according to mean grammaticality accuracy and reading times on the last word of the sentence.

Table 3.3 Descriptive statistics for statistical learning task

<table>
<thead>
<tr>
<th></th>
<th>Grammaticity accuracy (%)</th>
<th></th>
<th>Reading times on the final word (ms)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiar</td>
<td>Novel</td>
<td>Mean Score</td>
<td>Familiar</td>
</tr>
<tr>
<td>Grammatical</td>
<td>77</td>
<td>65</td>
<td>71</td>
<td>1941.35</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td>57</td>
<td>57</td>
<td>57</td>
<td>1947.68</td>
</tr>
<tr>
<td>Mean score</td>
<td>77</td>
<td>61</td>
<td>64</td>
<td>2315.24</td>
</tr>
</tbody>
</table>

The mean grammaticality accuracy (calculated from the percentage of correct responses) is 64% ($SD = 24\%$), with the scores ranging from 10 to 100: 71% ($SD = 24\%$) for grammatical trials and 57% ($SD = 32\%$) for ungrammatical trials. The one sample $t$-test shows that the mean grammaticality accuracy is significantly higher than chance level (i.e., the test value set at 50\%): $t(52) = 19.620, p = .0001$. As in Table 3.3, the grammaticality accuracy on grammatical trials is significantly higher than that on ungrammatical trials (71\% vs. 57\%) [$t(52) = 3.514, p = .001$]. On grammatical trials, the grammaticality accuracy of familiar sentences is significantly higher than the grammaticality accuracy of novel sentences (77\% vs. 65\%), [$t(52) = 3.884, p = .0001$]. Crucially, however, the mean grammaticality accuracy of novel sentences (65\%) is significantly higher than chance level [an one sample $t$-test, $t(52) = 16.984, p = .0001$]. This shows that participants are able to notice the key dependency even in sentences not previously encountered in the training session.
The mean reading time (calculated from reading times on the last word of the test sentences) was 2132.53ms ($SD = 1339.12$ms), with a range from 541.76 to 7768.18: for grammatical trials, the mean reading time was 2281.14ms ($SD = 1554.88$ms), and for ungrammatical trials, 1947.68ms ($SD = 1910.31$ms). A paired $t$-test showed no significant difference of reading times on the final word in grammatical and ungrammatical sentences [$t(51) = -0.059$, $p = .953$]. However, the reading times on the final word on grammatical trials is significantly slower in the familiar trials than in the novel trials [$t(51) = -2.620$, $p = .012$].

In summary, these results replicated results from previous studies on statistical learning and show accuracy rates on the grammaticality task similar to those of previous studies: for example, 69.2% for nonadjacent statistical learning from Misyak and Christiansen’s (2012) study compared to 64% in the current study.

3.2.2 Self-paced reading task: Processing English relative clauses

For the self-paced reading task, mean comprehension accuracy rates were calculated from the percent of correct responses. The mean comprehension accuracy is 80% ($SD = 17\%$) - 87% ($SD = 16\%$,) for SRs and 73% ($SD = 20\%$) for ORs, with 90% ($SD = 4\%$) for fillers. The paired $t$-test shows that the mean comprehension accuracy for SRs is significantly higher than that for ORs [$t(52)= 5.8181$, $p = .0001$]. This asymmetry for subject and object relative clauses indicates the relative difficulty of comprehending object relative clauses.

We turn now to the reading time results. Overall, reading times in both of the SR and OR conditions was almost identical, except for regions 6, 7, and 8, as seen in Figure 3.1.
At region 6 (the word before the main verb *wrote*), reading times in object relative clauses were significantly slower than reading times in subject relative clauses (according to a paired sample *t*-test, $t(52) = 3.079, p = .003$), indicating that participants immediately noticed the potential gap position. At region 7 (at the main verb *wrote*), reading times in ORs were significantly slower than reading times in SRs; a paired sample *t*-test, $t(52) = 2.598, p = .012$. At region 8 (the word after the main verb *wrote*), there was no significant difference between SRs and ORs. Results from reading times at these three regions indicate that participants have difficulty processing ORs, as reported in several L1 language acquisition and processing studies (in L1 processing, Gibson, 1998; Gordon, Hendrick, & Johnson, 2001; King & Just, 1991; Lewis et al., 2006; in L1 language acquisition, Booth, MacWhinney & Harasaki, 2000; Friedmann & Novogrodsky, 2007; Kidd, Brandt, Lieven & Tomasello, 2007). In other words, performance on ORs in this study confirms the results of previous studies that found difficulty in processing ORs.
3.2.3 Combined results: Correlation between SL and online comprehension tasks

This dissertation examines a possible correlation between the results of the SL task involving nonadjacent dependencies in the artificial language and a self-paced reading task involving English direct object relative clauses. It is assumed that the manner in which nonadjacent dependencies are processed in the artificial language is similar in relevant respects to the way in which they are processed in English direct object relative clauses. This section will focus on the comprehension accuracy and residual reading times for direct object relative clauses, because the findings of previous studies showed ceiling effects for subject relative clauses.

In data analyses for this experiment, grammaticality accuracy on the SL task was used as a between-subjects variable. The participants were divided into good and poor
statistical learners based on the cut-off point set at $90^{11}$. Sixteen participants who scored above or equal to 90 were categorized as good statistical learners, while thirty-seven participants who scored below 90 were categorized as poor statistical learners. The mean grammaticality accuracy for good statistical learners was $97.5 \ (SD = 3)$, with a range from 93 to 100. In contrast, the mean grammaticality accuracy for poor statistical learners was $50.19 \ (SD = 65)$, with a range from 10 to 75. The mean reading time on the last word for good statistical learners was $1935.80 \text{ms} \ (SD = 1280.29)$, while for poor statistical learners, it was $2254.30 \text{ms} \ (SD = 1369.32)$.

Mean comprehension accuracy rates for good and poor statistical learners, shown in Figure 3.2, were calculated for the OR conditions. For the OR condition, the mean comprehension accuracy rate of both good and poor statistical learners was $73\% \ (SD = 25$ and 19, respectively), indicating no significant difference between the two groups [$t(51) = -0.044, p = .965$].

---

11 I opted for this value because the traditional way of splitting the group by its median in previous L1 and L2 processing studies typically results in rather homogeneous groups, so that effects of the dividing variable are often not observed (for discussion, Hopp 2013; Conway et al., 2005).
The mean residual reading time was calculated for each of the SR and OR conditions, but only the mean residual reading time for the OR condition is considered here. Figure 3.3 shows the residual reading times profiles of good and poor statistical learners. Overall, reading times of ORs was longer than those of SRs at the critical regions 6, 7, and 8, indicating that both of the groups encountered relative difficulty in processing ORs (see Appendix J). For the OR condition, there was no significant difference at critical regions 7 and 8 between good and poor statistical learners $[t(51) = 1.603, p = .549, t(51) = 1.173, p = .246$, respectively], but there was a significant difference at region 6. Let us look in more detail at reading times at region 6. Figure 3.4 shows mean residual reading times at region 6 in the OR condition.
Figure 3.3 English native speakers’ mean residual RTs profile by good and poor SLers in the OR condition.
For region 6 in the OR condition, the mean residual reading time of good statistical learners was 132.64ms ($SD = 398.06$), and the mean residual reading time of poor statistical learners was 444.49ms ($SD = 625.67$): here, the independent t-test showed a significant difference between the two groups [$t(51) = -2.179, p = .032$]. This indicates that participants with high grammaticality accuracy scores read significantly faster at region 6 than participants with low grammaticality accuracy scores, who presumably experienced a greater integration cost at this position.

![Mean residual reading times at region 6 for English native group](image)

**Figure 3.4 Mean residual reading times at region 6 for English native group**

The first objective in this analysis is to determine the relationship between the SL task involving nonadjacent dependencies and the online comprehension task involving direct object relative clauses. I explore the possibility of a significant correlation between two
relevant factors from the SL task (i.e., grammaticality accuracy and reading times on the final word) and two measures from the self-paced reading task (i.e., comprehension accuracy as well as reading times at the critical region of ORs).

Table 3.4 Intercorrelations between self-paced reading task (top row) and statistical learning (left column) by English native speakers

<table>
<thead>
<tr>
<th>SL Task</th>
<th>Comprehension Accuracy</th>
<th>Reading Times at Region 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>RC</td>
</tr>
<tr>
<td>Grammaticality Accuracy</td>
<td>.076</td>
<td>.071</td>
</tr>
<tr>
<td>RT at the final word</td>
<td>-.133</td>
<td>-.176</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01 (two tailed, n=53).

A significant correlation between statistical learning ability and comprehension of English ORs was not observed ($r = .076, p > .05$), as in Table 3.4. However, a significant correlation between statistical learning ability and processing of English ORs was found. As shown in Table 3.4, mean grammaticality accuracy on the SL task is negatively correlated with the reading times at the critical region 6 of ORs ($r = -.27, p < .05$).

Moreover, I found a positive correlation between reading times on the final words of the SL task and reading times at the critical region of English ORs ($r = .38, p < .01$).

To determine how well two measures on the SL task (i.e., mean grammaticality accuracy and reading time on the final word) predict comprehension and reading times for English ORs, a simple linear-regression was conducted. As seen in Table 3.5, regression coefficients of simple linear regression showed that mean grammaticality accuracy on the SL task cannot predict comprehension accuracy for ORs, but that grammaticality accuracy on the SL task predicts speed of reading for ORs ($\beta = -.27, p <$
Moreover, the reading times on the final word in the SL task predict speed of reading at the vicinity of main verb in ORs ($\beta = .40, p < .05$).

<table>
<thead>
<tr>
<th>Processing English ORs</th>
<th>Grammaticality Accuracy</th>
<th>RT at the final word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension accuracy</td>
<td>.08</td>
<td>-.13</td>
</tr>
<tr>
<td>RT at the critical region 6</td>
<td>-.27*</td>
<td>.40*</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01 (one tailed, n=53).

### 3.3 Discussion

In summary, the results of Experiment 1 show, on the one hand, that the results from both the SL task and the English relative clause processing task replicated previous studies. First, results from the SL task demonstrate that the English native speakers were able to track a nonadjacent dependency in an artificial language: they achieved a mean grammaticality accuracy rate of 64%, which is similar to rates achieved in previous research (see Gómez, 2002; Misyak & Christiansen, 2012). Second, findings from the self-paced reading task confirm the difficulty of processing ORs, as described in the L1 processing literature (Gibson, 1998; Gordon, Hendrick, & Johnson, 2001; King & Just, 1991; Lewis et al., 2006): the current study shows (1) that the comprehension accuracy rate in the SR condition was significantly higher than that in the OR condition, and (2) that the English native speakers encountered difficulty with ORs, as shown by the longer RTs at the vicinity of main verb (Figure 3.1). On the other hand, the present study found an effect of statistical learning ability on the processing of English object relative clauses. Let us discuss these findings in turn.
First, even though I did not find an effect of statistical learning ability on the comprehension of English object relative clauses, I did find a significant correlation between statistical learning ability and the processing of English ORs. A simple linear regression showed that mean grammaticality accuracy predicts the speed of reading time in ORs, showing that good statistical learners read significantly faster at the critical region (the vicinity of main verb) than poor statistical learners did. Moreover, the reading times on the final word of the SL task predict the speed of reading times at the critical region as in Table 3.5. These findings replicated the results from Misyak, Christiansen, & Tomblin’s (2010) study, in which Cornell undergraduates with higher nonadjacent statistical learning scores read significantly faster than those with lower nonadjacent statistical learning scores at the vicinity of main verb in their native English.

Second, the degree of statistical learning ability is likely to explain individual differences in English native speakers’ performance. Several studies have tried to explain individual variation in language comprehension by focusing on the role of working memory in L1 processing (King & Just, 1991; Just & Carpenter, 1992). Assuming that statistical learning ability is considered a kind of cognitive ability, the findings of the current study provide additional evidence to explain individual differences in online comprehension. In the capacity-based perspective, poor statistical learners’ poor processing of ORs reflects a limitation in memory resources, which leads to individual variation even in English native speakers’ comprehension as outlined in previous studies on the relation between working memory and online comprehension. As shown in Figure 3.5, variation in grammatical accuracy rates on the SL task appears to predict English native speakers’ speed of reading ORs, but to test this possibility, it is necessary to test a
different group.

In addition, no effect of statistical learning ability on comprehension accuracy was observed—even though Misyak and Christiansen (2012) found a correlation between nonadjacent dependency learning ability and a total comprehension accuracy for English relative clauses. However, their study did not include an attempt to examine between grammaticality accuracy in the SL task and comprehension accuracy in ORs distinct for SRs, as reported in my experiment. It is worth noting that while ORs have a non-trivial filler-gap dependency according to standard analyses, simple SRs do not, as no referents intervene between the filler and the gap.

The comparison between statistical learning and online comprehension sheds light on the effect of statistical learning ability, in that the different levels of statistical learning
ability predict individual variation in L1 processing. Statistical learning ability is, of course, not the only factor involved in individual variation in language processing and acquisition, but to understand its role in processing and acquisition, it is necessary to test different groups and speakers of different languages. For this purpose, I designed and conducted Experiment 2.
CHAPTER 4
EXPERIMENT 2: KOREAN LEARNERS OF ENGLISH

This chapter presents an experiment consisting of two tasks that were completed by Korean learners of English. The experiment was designed to investigate the role of statistical learning ability in second language (L2) sentence processing and comprehension. This question has never been tested in the context of L2 learning. To test second language learners’ (L2ers’) sensitivity to the processing of nonadjacent dependencies, the same two tasks employed in Experiment 1 with English native speakers were conducted with Korean learners of English. In comparing the results from these two tasks, I hope to shed light on the role of statistical learning ability in explaining individual differences in L2 processing.

This experiment with Korean learners of English was designed to directly test the effect of statistical learning ability on the processing of object relative clauses. The research questions addressed in this chapter are as follows:

a. Can Korean learners of English, like the English native speakers in Experiment 1, learn from sequentially presented input a simple grammar that includes a nonadjacent dependency similar to the long-distance dependency in English object relative clauses?

b. Is an effect of L2ers’ proficiency on a C-test observed in comprehension accuracy rates and in the reading times at the critical regions of object relative clauses?
c. Is an effect of statistical learning ability observed in the comprehension of English RCs? It is predicted that higher grammaticality accuracy rates in the statistical learning task will correlate with higher comprehension accuracy rates in processing English object relative clauses and with faster reading times at the critical regions.

d. Is an effect of English proficiency observed on the relation between statistical learning ability and the processing of object relative clauses?

4.1 Method

4.1.1 Participants

A total of 114 Korean learners of English as a second language (87 women and 27 men; mean age = 21.3, SD = 1.7, range = 19–26), all of them Korea University undergraduates, participated for course credit or money. Most of them were majoring in English, English education, linguistics, and other disciplines; their TOEIC scores ranged from 735 to 990. Like the English native speakers in Experiment 1, they all completed two tasks: a statistical learning task and a self-paced reading task. Table 4.1 below summarizes the participants’ mean length of residence in English-speaking environments (e.g., US, UK, Canada, Australia) and the participants’ mean length of exposure to English grammar, as well as their scores on an English C-test (max = 40) adapted from Schulz’s (2006) dissertation (see Appendix I). All of the information was gathered as a part of Experiment 2, reported below. The background questionnaire is given in Appendix H.
Table 4.1 Profiles of the Korean learners of English

<table>
<thead>
<tr>
<th>Groups</th>
<th>M:F ratio (N=</th>
<th>Mean Age (max=40)</th>
<th>C-test (max=40)</th>
<th>Length of residence (years)</th>
<th>Length of exposure (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>14:51 (65)</td>
<td>21.2 (1.6)</td>
<td>30.6 (2.6)</td>
<td>0.9 (2)</td>
<td>11.49 (2.5)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>13:36 (49)</td>
<td>21.5 (1.8)</td>
<td>21.6 (3.4)</td>
<td>0.2 (0.6)</td>
<td>10.35 (2.9)</td>
</tr>
<tr>
<td>Total</td>
<td>27:87 (114)</td>
<td>21.3 (1.7)</td>
<td>26.7 (5.4)</td>
<td>0.6 (1.6)</td>
<td>11 (2.4)</td>
</tr>
</tbody>
</table>

Note. Standard deviation is given in parentheses.

As Table 4.1 shows, the total mean score on the English C-test was 26.7 (SD = 5.4), which is over 66% of the full score of 40. For advanced L2ers, the mean score on the C-test was 30.6 (SD = 2.6), which is over 76.5% of the total score of 40. For intermediate L2ers, the mean score on the C-test was 21.6 (SD = 3.4), which is over 53% of the full score of 40. An independent $t$-test showed that advanced L2ers’ English proficiency on the C-test was significantly higher than that of intermediate L2ers [$t(112) = 15.536, p = .0001$]. These results indicate that the advanced L2ers were significantly advanced compared to intermediate L2ers, so an effect of C-test proficiency on both tasks was expected. Moreover, there were significant differences between advanced and intermediate groups in both mean length of residence in English-speaking countries and mean length of exposure to English grammar [$t(112) = 2.266, p = .025$, $t(112) = 2.858, p = .005$].

4.1.2. Materials

The materials used in Experiment 2 were identical to those described in the previous chapter.
4.1.3 Procedure

The procedure was also identical to that of Experiment 1, except for the C-test.\footnote{A C-test is very similar to a Cloze test. The main difference between a C-test and a Cloze test is that the first half of a word is provided so that no complete words are missing. The C-test used for this experiment consisted of 40 items, each one comprising half of a word plus a blank. The highest possible score was 40. Participants who scored from 40 to 27 were classified as advanced, and those who scored from 27 to 16 were classified as intermediate.} The C-test and language background questionnaire were conducted with each participant immediately after he or she completed the two tasks. For the C-test, participants were informed that they would be shown the first half of a number of words and that they would be given 10 minutes to complete all the words. Following this, they were to fill in the answers to the language background questionnaire. The whole experimental session took 45 minutes to one hour for each participant; the L2ers generally took longer than the English native speakers who participated in Experiment 1.

4.1.4 Data analysis

The data analysis method for this experiment was identical to that of Experiment 1, described in section 3.1.4. To examine the effect of English proficiency on the relation between statistical learning ability and the processing of object relative clauses, an additional analysis was conducted by splitting participants into advanced and intermediate groups based on the cut-off value of 27 on the C-test.

4.2 Results

This section reports results from the Korean learners of English. First, I provide a descriptive analysis of the results from the two tasks. I then present the relation between
the effect of statistical learning ability on the processing of object relative clauses, on the one hand, with the effect of English proficiency, on the other.

4.2.1 Statistical learning task

Table 4.2 provides the descriptive statistics for the L2ers’ performance on the statistical learning (SL) task according to mean grammaticality accuracy rate and reading time on the final word of each sentence.

<table>
<thead>
<tr>
<th></th>
<th>Grammaticality accuracy (%)</th>
<th>Mean Score</th>
<th>Reading times on the final word (ms)</th>
<th>Mean RT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grammatical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Familiar</td>
<td>80</td>
<td>71</td>
<td>1459.12</td>
</tr>
<tr>
<td></td>
<td>Novel</td>
<td>71</td>
<td>75</td>
<td>1577.76</td>
</tr>
<tr>
<td></td>
<td>Ungrammatical</td>
<td>61</td>
<td>61</td>
<td>1577.76</td>
</tr>
<tr>
<td>Mean score</td>
<td>80</td>
<td>66</td>
<td>68</td>
<td>1459.12</td>
</tr>
</tbody>
</table>

The mean grammaticality accuracy was 68% ($SD = 23$), with the scores ranging from 28 to 100. The mean accuracy rates for grammatical and ungrammatical trials were 75% ($SD = 23$) and 61% ($SD = 32$), respectively. A one sample $t$-test showed that the total mean grammaticality accuracy rates were significantly higher than chance level (i.e., the test value set at 50%) [$t(113) = 31.761, p = .0001$]. This replicates the results of previous studies that found that English native speakers were able to extract rules from sequential input, but no previous study has tested this statistical learning ability with L2ers.

In sum, as expected, grammaticality accuracy scores on grammatical trials are significantly higher than that on ungrammatical trials (75% vs. 61%) [$t(113) = 5.067, p = .0001$]. Moreover, on grammatical trials, the grammaticality accuracy rate is
significantly higher for familiar sentences than for novel sentences (80% vs. 71%), \([t(113) = 4.741, \ p = .0001]\). Crucially, however, the mean grammaticality accuracy rate is significantly higher than chance for novel sentences (65%) [a one sample \(t\)-test, \(t(113) = 27.111, \ p = .0001\)]. This suggests that participants are able to notice the key dependency even in sentences not presented in a training session.

Mean reading times were calculated from reading times on the final word of test sentences. The mean reading time across all conditions was 1519.95ms \((SD = 1063.32ms)\), with reading times ranging from 330.04 to 6538.26. The mean reading times for grammatical and ungrammatical trials were 1473.03ms \((SD = 995.31ms)\) and 1577.76ms \((SD = 1362.73ms)\), respectively. A paired \(t\)-test showed that reading times on the final word were significantly faster for grammatical sentences than for ungrammatical sentences \([t(110) = -3.129, \ p = .002]\). On grammatical trials, no significant difference on reading times between familiar and novel sentences was observed \([t(112) = -0.086, \ p = .932]\).

This experiment was also designed to test the relation between statistical learning ability and L2ers’ English proficiency. No effect of proficiency was observed on either grammaticality accuracy or reading times in the SL task \([t(112) = 1.191, \ p = .236, \ t(112)= .203, \ p = .839, \text{ respectively}]\). In summary, like the English native speakers in Experiment 1, Korean native speakers (who happen also to be second language learners) were able to identify and track a nonadjacent dependency.
4.2.2 Self-paced reading task: Processing English relative clauses

For the self-paced reading task, the mean comprehension accuracy rate by all L2ers across conditions was 81% ($SD = 17$). The mean rates for the SR and OR conditions were 83% ($SD = 17$) and 79% ($SD = 20$), respectively, with a mean rate of 88% ($SD = 6$) for fillers. This asymmetry between subject and object relative clauses indicates the significant difficulty of processing object relative clauses [a paired sample $t$-test, $t(113)=2.296, p = .024$]. For the advanced L2 group, the mean comprehension accuracy rate was 85% ($SD = 13$): 86% ($SD = 15$) for SRs and 81% ($SD = 16$) for ORs, with 88% ($SD = 6$) for fillers. For the intermediate L2 group, the mean comprehension accuracy rate was 77% ($SD = 19$): 84% ($SD = 17$) for SRs and 74% ($SD = 23$) for ORs, with 87% ($SD = 8$) for fillers. A significant difference between subject and object relative clause is observed among intermediate L2 group, but not in the advanced L2 group.

![Figure 4.1 Mean comprehension accuracy by advanced and intermediate L2ers](image-url)
A repeated-measures ANOVA with C-test proficiency as a between-subjects variable and the type of relative clause (SR vs. OR) as a within-subjects variable showed a main effect of relative clause type \([F(1,112) = 6.127, p = .015]\), indicating that the processing of object relative clauses was significantly more difficult than that of subject relative clauses. A main effect of C-test proficiency was also observed \([F(1,112) = 6.381, p = .013]\), indicating that the advanced L2ers comprehended relative clauses significantly better than the intermediate L2ers on both relative clause types. There was no significant interaction between relative clause type and proficiency level \([F(1,112) = 1.890, p = .172]\), indicating that different accuracy rates between the SRs and ORs was not due to different proficiency levels. Even though there was no significant interaction between relative clause type and proficiency level, results confirmed the findings of previous studies on the difficulty of processing ORs. Furthermore, the advanced L2 group showed higher comprehension accuracy in ORs than the intermediate L2 group did (81% vs. 74%) [an independent t-test, \(t(112) = 2.579, p = .012\)], whereas advanced and intermediate L2 groups did equally well on SRs in terms of comprehension accuracy [an independent t-test, \(t(112)=1.581, p = .118\)]

Figure 4.2 shows L2ers’ reading times across all regions of sentences. Mean residual reading times were calculated from all words in the test sentences. Regions 6, 7, and 8 are of particular interest because this is where the reader is likely to slow down to keep track of the dependency. For the advanced L2ers, at region 6 (the word before the main verb *wrote*) and region 7 (the main verb), residual reading times for ORs were significantly slower than reading times for SRs [a paired sample t-test, \(t(64)= 4.430, p = .0001, t(64) =2.850, p = .006\), respectively]; at region 8 (the word after the main verb *wrote*), there
was no significant difference between SRs and ORs \(t(64) = .473, p = .638\). For the intermediate L2ers, at regions 6 and 7, residual reading times for ORs were significantly slower than reading times for SRs [a paired sample \(t\)-test, \(t(48) = 2.213, p = .032, t(48) = 2.077, p = .043\), respectively]; at region 8, there was no significant difference between SRs and ORs \(t(48) = .106, p = .916\). Both L2 groups, then, showed an asymmetry between subject and object relative clauses in regions 6 and 7, indicating greater processing difficulty for ORs at those two points.

Let us look at the effect of L2 proficiency on the processing of English relative clauses. For the SR condition, despite numerical differences, no significant difference between the advanced and intermediate L2 groups was observed across critical regions (6, 7, and 8). As shown in Figure 4.2, the same was true for the OR condition: no significant differences between the advanced and intermediate L2 groups were observed across critical regions 6, 7, and 8 \(t(112) = .282, p = .779, t(112) = .652, p = .516, t(112) = -.200, p = .842\), respectively]. However, although I found no effect of proficiency on reading speed at critical regions 6 and 7 in ORs, the comprehension accuracy rates discussed above indicate that all L2ers, regardless of proficiency, found ORs more difficult to process than SRs.
Figure 4.2 Korean L2ers’ mean residual RT’s profile for English RCs
4.2.3 Combined results: Correlation between SL and online comprehension tasks

This dissertation aims to look at the possible correlation of performance on the SL task involving nonadjacent dependencies in an artificial language and performance on a self-paced reading task involving English direct object relative clauses. This section will focus on the comprehension accuracy and residual reading times of direct object relative clauses.

In this section, I examine the question of the role of statistical learning ability on the processing of ORs by using grammaticality accuracy rate on the SL task as a between-subjects variable. Like the English native speakers in Experiment 1, all L2ers were divided into good statistical learners and poor statistical learners based on the cut-off point of 90. The 37 participants who scored above or equal to 90 were categorized as good statistical learners, while the 77 participants who scored below 90 were categorized as poor statistical learners. The mean grammaticality accuracy rate for good statistical learners was 96 ($SD = 3$), with a range of 90 to 100. The mean grammaticality score for the poor statistical learners was 55 ($SD = 13$), with a range of 28 to 88. Mean reading times on the last word of the sentence for the good and the poor statistical learners were 1263.85ms ($SD = 525.17$) and 1828.89ms ($SD = 1202.28$), respectively. To determine whether a correlation between grammaticality accuracy and reading time on the final word existed, a simple linear-regression was conducted. It showed that the higher the grammaticality accuracy was, the faster the reading times were on the final word [$t(112) = -3.968, p = .0001$].

For the self-paced reading task, mean comprehension accuracy rates were calculated for the SR and OR conditions, but this section will focus on the results from the OR
condition. As shown in Figure 4.3, for the OR condition, the mean comprehension accuracy rates of good and poor statistical learners were 91% ($SD = 13\%$) and 75% ($SD = 21\%$), respectively, with a significant difference between the two groups [$t(112) = 4.966, p = .0001$]. As expected, good statistical learners did better than poor statistical learners in comprehending ORs.

![Figure 4.3 Mean comprehension accuracy on ORs for all L2ers](image)

Figure 4.3 Mean comprehension accuracy on ORs for all L2ers

Figure 4.4 shows mean residual reading time (RT) profiles for all L2ers on object relative clauses. (Mean residual RTs were calculated for all words of the sentences.) Overall, all L2ers read SRs faster than ORs across all regions, indicating that more memory resources are needed for processing ORs (see Appendix J). The critical regions of interest are 6, 7, and 8 in ORs, where higher processing load is expected.
Figure 4.4 All L2ers’ mean residual RTs profile by good and poor SLers in English ORs
As in Figure 4.4, for the OR condition, there was no significant difference between good and poor statistical learners at regions 6 and 7 \( t(112) = 1.260, p = .260, t(112) = 1.504, p = .137 \), respectively. However, as shown in Figure 4.5, at region 8 (a spillover position, after the main verb *wrote*), good statistical learners unexpectedly read significantly slower than poor statistical learners [an independent \( t \)-test, \( t(112)= 2.199, p = .03 \)]. This might indicate that good statistical learners read more carefully, which results in relatively higher comprehension accuracy compared to poor statistical learners (see Figure 4.3). This possibility will be explored in more detail in the general discussion below.

![Figure 4.5 Mean residual reading times at region 8 for all L2ers in ORs](image)

Figure 4.5 Mean residual reading times at region 8 for all L2ers in ORs
4.2.4 The effect of L2ers’ English proficiency

To examine the interaction of comprehension accuracy and reading times with L2ers’ English proficiency, each of the good and poor statistical learners were placed into one of two groups (advanced L2 group vs. intermediate L2 group) based on the cut-off value of the C-test (27 out of 40). The 65 participants (21 good SLers and 44 poor SLers) who scored above or equal to 27 were categorized as advanced L2ers, whereas the 49 participants (16 good SLers and 33 poor SLers) who scored below 27 were categorized as intermediate L2ers.

Figure 4.6 displays the mean comprehension accuracy rates of the good and poor statistical learners in the advanced L2 group.

Figure 4.6 Mean comprehension accuracy on English ORs for advanced L2 group
For the OR condition, the mean comprehension accuracy of good and poor statistical learners was 90\% (SD = 14) and 81\% (SD = 17), respectively, with a significant difference between the two groups \([t(63) = 2.405, p = .020]\). This demonstrates that sensitivity to nonadjacent statistical learning significantly affected comprehension in the OR condition.

Next, let’s look at the reading times (RTs) at critical region 8. Figure 4.7 shows the mean residual RTs at region 8 of the good and poor statistical learners in the advanced L2 group. In the OR condition, however, at region 8, good statistical learners unexpectedly read significantly slower than poor statistical learners [an independent \(t\)-test, \(t(63) = 2.373, p = .025\)]. As noted earlier, this may have been because good statistical learners read more carefully, which led to relatively high comprehension scores (see Figure 4.6). This finding is different from that for English native speakers in Experiment 1: good statistical learners read much faster than poor statistical learners at the critical region 6.

![Figure 4.7 Mean residual reading times at region 8 for advanced L2 group in ORs](image)
Now, let us look at results from the intermediate L2 group. As seen in Figure 4.8, for the OR condition, the mean comprehension accuracy rates of good and poor statistical learners was 91% ($SD = 12$) and 66% ($SD = 23$), respectively, with a significant difference between the two groups [$t(47) = 4.840, p = .0001$]. As in the advanced L2 group, good statistical learners significantly outperformed poor statistical learners.

![Figure 4.8 Mean comprehension accuracy on English ORs for intermediate L2 group](image)

Figure 4.8 Mean comprehension accuracy on English ORs for intermediate L2 group

Figure 4.9 shows the mean residual RTs at critical region 8 for good and poor statistical learners in the intermediate L2 group. In the OR condition, the RTs at region 8 (a spillover position, after the main verb *wrote*) for good and poor statistical learners were almost identical [an independent $t$-test, $t(47) = .596, p = .545$]. Thus, no significant difference between good and poor statistical learners is observed in OR condition.
Drawing on data from two tasks involving 114 L2ers, the first objective in this analysis is to determine whether there is a correlation between performance on the SL task involving nonadjacent dependencies and performance on the online comprehension task involving direct relative clauses. Thus, I search for a significant correlation between two factors derived from the SL task (i.e., grammaticality accuracy and reading times on the final word) and two measures from the self-paced reading task (i.e., comprehension accuracy as well as reading times at the critical region of ORs).

Table 4.3 Intercorrelations between self-paced reading task (top row) and statistical learning (left column) with C-test scores by L2ers

<table>
<thead>
<tr>
<th></th>
<th>Comprehension Accuracy</th>
<th>Reading Times at Region 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>RC</td>
</tr>
<tr>
<td>Grammaticality accuracy</td>
<td>.324**</td>
<td>.289*</td>
</tr>
<tr>
<td>RT at the final word</td>
<td>-.071</td>
<td>-.078</td>
</tr>
<tr>
<td>Proficiency (C-test scores)</td>
<td>.238*</td>
<td>.253*</td>
</tr>
</tbody>
</table>

†p < .10, *p < .05, **p < .01 (two tailed, n=114).
A significant correlation between statistical learning ability and comprehension of English ORs was observed ($r = .324, p < .05$), as in Table 4.3. In addition, a significant correlation between statistical learning ability and comprehension of English RCs was found ($r = .289, p < .05$). I also found a marginally significant correlation between mean grammaticality accuracy on the SL task and reading times at the critical region 8 in ORs ($r = .157, p < .09$). Finally, turning to L2ers’ English proficiency, I found that C-test scores is positively correlated with the comprehension of English ORs.

To determine how well mean grammaticality accuracy on the SL task predicted language comprehension, a simple linear-regression was conducted. Even though I did not find an effect of statistical learning ability on processing of English ORs, I did find a significant correlation between statistical learning ability and comprehension of English ORs. As seen in Table 4.4, the regression coefficients of simple linear regression showed that mean grammaticality accuracy predicts the comprehension in ORs ($\beta = .324, p < .05$). However, the reading times on the final word of the SL task does not predict the speed of reading times at the critical region ($\beta = -.71, p > .05$). Moreover, C-test scores, as a measure of relative English proficiency, are positively correlated with the comprehension of English ORs ($\beta = .24, p < .05$).

<table>
<thead>
<tr>
<th>Processing English ORs</th>
<th>Statistical learning task</th>
<th>Proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grammaticality accuracy</td>
<td>RT at the final word</td>
</tr>
<tr>
<td>Comprehension accuracy</td>
<td>.32**</td>
<td>-.71</td>
</tr>
<tr>
<td>RT at the critical region 8</td>
<td>.16</td>
<td>-.15</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01 (one tailed, n=114).
4.3 Discussion

To summarize, the results of Experiment 2 show that Korean native speakers (who happen also to be L2 learners of English) are able to track nonadjacent dependencies in an artificial language presented sequentially, just as English native speakers did in Experiment 1: the grammaticality accuracy rate for the English native speakers was 64%, compared to 68% for Korean native speakers. These rates confirm the results from previous studies on statistical learning by showing that L2ers are also sensitive to the statistical cues that are essential to learning nonadjacent dependencies in the SL task. But does this have an effect on second language learning?

As with the statistical learning task, results from processing English relative clauses also confirm results as outlined in previous studies, in that object relative clauses were found to be more difficult than subject relative clauses: that is, with ORs, participants showed lower comprehension accuracy rates and slower reading times at the critical regions of the OR. An effect of English proficiency on RCs was observed only in the OR condition: as Figure 4.1 shows, there was a significant difference between advanced and intermediate L2 groups (81% vs. 74%) in comprehension rates in the OR condition but not in the SR condition. It seems, then, that neither intermediate nor advanced L2ers have difficulty processing SRs.

One of the research questions that the present study was designed to address is whether L2ers’ statistical learning ability affects their performance in processing object relative clauses. Because the SL task and the processing of ORs both involve nonadjacent dependencies, the prediction is that good statistical learners will perform better on OR
comprehension than poor statistical learners, with faster reading times at the critical regions. Let us discuss the findings in turn.

First, based on data from 114 L2ers, the effect of nonadjacent statistical learning on processing OR was observed in that (1) good statistical learners comprehended ORs significantly better than poor statistical learners did and (2) there was significant difference in reading times at region 8 between good and poor statistical learners. However, unlike good statistical learners within the English native speaker group, good statistical learners within the L2 group tended to read significantly slower than poor statistical learners did. This study also tested the effect of L2ers’ English proficiency on the comprehension of relative clauses by splitting participants into advanced and intermediate L2 groups. In the advanced L2 group, good statistical learners did better than poor statistical learners in comprehending ORs, and they had slower reading times than poor statistical learners at critical region 8, as seen in all good statistical learners (see Figure 4.5). The slower reading correlates with higher comprehension accuracy in the OR condition, but it is unclear why. It is perhaps worth considering the possibility that good statistical learners in the L2 group attend more carefully to the dependency between the filler and its gap position, thereby slowing their reading time. In the intermediate L2 group, good statistical learners performed better than poor statistical learners in comprehending ORs, but no difference in reading times in the OR condition was observed between good and poor statistical learners.

I summarize the findings from Experiment 2 based on Table 4.3. It shows that mean grammaticality accuracy in the SL task is correlated with both the comprehension accuracy rate of object relative clauses ($r = .324, p < .05$) and the total comprehension
accuracy rate of the two relative clauses ($r = .289$, $p < .05$), as predicted. This correlation is also observed in Misyak & Christiansen (2012), but their study did not report a direct correlation between nonadjacent statistical learning and comprehension accuracy in the OR condition. Moreover, mean grammaticality accuracy is likely to predict the speed of reading times at critical region 8, ($r = .157$, $p < .10$). This marginally positive correlation indicates that good statistical learners read more slowly than poor statistical learners at this region. However, reading times in the SL task do not predict performance in the self-paced reading task.

Overall, findings from Experiment 2 are not directly comparable with results from Misyak & Christiansen’s (2012) study because their study did not test L2ers’ performance. However, variation in the SL task is likely to predict not only L2ers’ comprehension accuracy for ORs but also L2ers’ reading times in the OR condition. In this regard, SL ability is likely to predict L2ers’ performance, especially their comprehension accuracy in the OR condition. Previous studies have failed to correlate individual differences in L2 online comprehension with L2ers’ working memory, and the role of working memory in L2 sentence processing has not been systematically examined across different structures. Perhaps, an explanation for individual differences will emerge through a deeper examination of statistical learning ability, to the extent that this ability reflects learners’ ability to extract rules from input.
CHAPTER 5

EXPERIMENT 3: KOREAN NATIVE SPEAKERS

This chapter presents an experiment consisting of two tasks that were completed by Korean native speakers. The experiment was designed to test the correlation between statistical learning ability and the processing of Korean filler-gap constructions. As far as I know, this is the first time that a correlation between statistical learning and sentence processing has been examined for a language other than English. Findings from Experiment 3 will contribute to understanding the role of statistical learning ability in L2 learning. As a further test of Korean native speakers’ learning of a nonadjacent dependency from sequentially exposed input, the results from the SL task in Experiment 2 were used for the purposes of comparison, so only Korean online comprehension task was conducted.

This experiment has a within-subject design in order to test both L1 influence on L2 processing and the effect of statistical learning ability on Korean filler-gap constructions. The research questions addressed in this chapter are as follows:

a. Do Korean native speakers show difficulty in processing Korean topicalized object sentences compared to topicalized subject sentences, as shown by comprehension accuracy and by reading times at critical regions?

b. Is an effect of statistical learning ability observed in processing Korean topicalized object sentences?

c. Is L1 influence observed in L2ers’ processing of English relative clauses, as shown by similar processing behavior?
5.1. Method

5.1.1 Participants

The same 114 Korean native speakers (87 women and 27 men; mean age = 21.3, SD = 1.7, range = 19-26) participated in Experiment 3 as in Experiment 2.

5.1.2 Materials

The SL task used in Experiment 3 was the same task described in chapter 3 (see section 3.1.2.1). This section thus describes only the materials of the Korean self-paced reading task. Although English relative clauses were used for the English self-paced reading task in Experiment 1 and 2, Korean relative clauses were not employed in Experiment 3 as the syntactic properties of English and Korean relative clauses are not comparable for the purposes of this experiment. As shown in (1b), English has postnominal relative clauses, in which the head noun boy (the filler) is associated with the gap position marked with an underscore; the relationship between the filler and the corresponding gap is established at a distance. The distance between the two positions in the OR imposes a heavy burden on working memory, compared to the relatively shorter distance between the two positions in the subject relative clause in (1a).

(1) English relative clause

a. SR: The boy*/that saw a girl*

b. OR: The boy/*that a girl saw */

In contrast, Korean has prenominal relative clauses without overt complementizers; the predicate of a Korean relative clause bears the relative clause verbal suffix -(nu)n, the so-called relativizer (O’Grady et al., 2003; 436), as in (2).
(2) Korean relative clause

a. [sonye-lul po-nun] sonyen,
girl—ACC see—RC.PRS boy
‘The boy that saw a girl’

b. [sonye-ka po-nun] sonyen/
girl-NOM see—RC.PRS boy
‘The boy that a girl saw’

In Korean ORs, as in (2b), the distance between a gap (underscore) and its filler (sonyen ‘a boy’) is relatively short compared to the gap-filler distance in Korean SRs, as in (2a).

In fact, Korean relative clauses contain not a filler-gap dependency, as in English relative clauses, but a gap-filler dependency, in that the processor encounters the gap (underscore) before its filler (sonyen ‘a boy’). This difference between English and Korean relative clauses means that the processing of RCs in these two languages must also be different, so I decided to choose an alternative Korean construction for this experiment, one that is structurally more similar to the filler-gap dependency that exists in English relative clauses.

I used Korean sentences with a sentential adverb, which canonically appears in sentence-initial position. Sentential adverbs are classified as structurally higher adverbs, following Cinque’s (1999) relative order of adverb phrases (Lee, 2000). Example (3) shows the canonical order of sentences that include Korean sentential adverbs (e.g., tahaynghi ‘fortunately’).

(3) Tahaynghi choykyoswu-nun ecey samwusil-eyse haksayng-ul cal mannass-ta.
fortunately choi profesor-TOP yesterday office-LOC student-ACC well met-DECL.
'Fortunately, Professor Choi met a student in his office yesterday.'
I used 8 Korean sentential adverbs: *tahaynghi* ‘fortunately’, *pwulhaynghi* ‘unfortunately’, *amwuthun* ‘anyway’, *solcikhi* ‘frankly’, *icey* ‘now’, *hoksi* ‘perhaps’, *thukhi* ‘particularly’, *pothong* ‘usually’. In the *Sejong* written and spoken corpus (a total of 10,697,678 words), these 8 adverbs appeared sentence-initially at least 75% of the time.

For this experiment, I manipulated sentences in two ways: by marking a subject with –*(n)*un (a topic marker), as in (4a), and by scrambling the positions of subject and direct object across the sentential adverb, as in (4).

(4) Korean topicalized sentences

Topicalized subject sentences
a. choykyoswu-nun, *tahaynghi* ___ ecey samwusil-eyse haksayng-ul cal choi profesor-TOP fortunately yesterday office-LOC student-ACC well mannass-ta. met-DECL.

'Fortunately, Professor Choi met a student in his office yesterday.'

Topicalized object sentences
b. haksayng-ul, *tahaynghi* choykyoswu-nun ecey samwusil-eyse_____j cal student-ACC fortunately choi profesor-TOP yesterday office-LOC well mannass-ta. met-DECL.

'Fortunately, Professor Choi met a student in his office yesterday.'

I marked the topicalized subject with –*(n)*un rather than –*ka/-i* (nominative markers), as the former option is more natural in a case where the subject is fronted.

Topicalized word order creates a filler-gap dependency similar to that of English relative clauses. That is, the relative difficulty of processing the two types of Korean sentences created through scrambling is similar to the relative difficulty of processing English SRs and ORs. In (4a), a subject (*choykyoswu-nun* ‘Professor Choi’) is fronted
across the sentential adverb (*tahaynghi* ‘fortunately’) from its original position. Similarly, in (4b), a direct object is fronted across both the sentence adverb (*tahaynghi* ‘fortunately’) and the subject (*choykoswu-nun* ‘Professor Choi’). As with English relative clauses, topicalized direct object sentences should be more difficult to comprehend than topicalized subject sentences because of the relatively greater distance between the position of the fronted element (the filler) and its original position (the gap), arguably imposing a greater processing load and higher memory cost (Nakayama, 1995; Mazuka, Itoh, and Kondo, 1997; Miyamoto and Takahashi, 2002). It thus seems reasonable to suppose that processing considerations are responsible for the contrast between the two patterns, consistent with the widely accepted assumption (e.g., Caplan & Waters, 2001, King & Just 1991, Carpenter, Miyake, & Just, 1994).

Experiment 3 involved 16 sets (2 sentences each) of experimental stimuli as illustrated in Table 5.1. Each stimulus had 7 regions. The critical regions were 4, 5, and 6, where participants’ reading times were expected to slow down in topicalized direct object sentences. The 16 sets of experimental stimuli were distributed across two lists. Each list comprised 16 target sentences and 32 filler sentences. In total, participants read 48 sentences. All fillers are complex sentences (see Appendix F).

Table 5.1 Sample set of experimental stimuli in Korean long-distance dependency experiment

<table>
<thead>
<tr>
<th>Regions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Topicalized subject sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choykoswu-nun tahaynghi ec ey samwusil-eyse haksayng-ul cal mannass-ta.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choi profesor-TOP fortunately yesterday office-LOC student-ACC well met-DECL.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Fortunately, Professor Choi met a student in his office yesterday.’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Topicalized direct object sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haksayng-ul tahaynghi choykoswu-nun ec ey samwusil-eyse cal mannass-ta.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>student-ACC fortunately Choi profesor-TOP yesterday office-LOC well met-DECL.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Fortunately, Professor Choi met a student in his office yesterday.’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In a self-paced reading task, participants were asked to read as naturally as possible and then to answer a yes/no comprehension question. All reading times for each region were recorded, along with responses to comprehension questions.

5.1.3 Procedure

Each of the 114 Korean learners of English completed a statistical learning task and an English relative clause processing task during the first session, after which they were asked to participate in a Korean self-paced reading task. This latter task was completed on a separate day from the first two tasks, with at least two days separating the sessions. The second session, involving the Korean self-paced reading task, took approximately 20 minutes for each participant.

5.1.4 Data analysis

The data analysis method for this experiment was identical to that of Experiment 1, described in section 3.1.4.

5.2 Results

5.2.1 Statistical learning

Due to a within-subject design, the descriptive statistics for Korean native speakers’ performance in the SL task is identical to those described in Experiment 2. The summarized results of the SL task, as in Table 5.2, are reproduced here.
Table 5.2 Descriptive statistics for statistical learning task by Korean native speakers

<table>
<thead>
<tr>
<th></th>
<th>Grammar (accuracy (%)</th>
<th>Mean Score</th>
<th>Reading times on the final word (ms)</th>
<th>Mean RT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiar</td>
<td>Novel</td>
<td>Familiar</td>
<td>Novel</td>
</tr>
<tr>
<td>Grammatical</td>
<td>80</td>
<td>71</td>
<td>75</td>
<td>1459.12</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td>61</td>
<td>61</td>
<td>61</td>
<td>1577.76</td>
</tr>
<tr>
<td>Mean score</td>
<td>80</td>
<td>66</td>
<td>68</td>
<td>1459.12</td>
</tr>
</tbody>
</table>

In sum, the mean grammaticality accuracy rate was 68% ($SD = 23$), and the scores ranged from 28 to 100; for grammatical and ungrammatical trials, the mean grammatical accuracy rates were 75% ($SD = 23$) and 61% ($SD = 32$), respectively. The one sample $t$-test shows that the mean grammaticality accuracy is significantly higher than chance-level (i.e., the test value set at 50%) [$t(113) = 31.761, p = .0001$]. Thus, as reported in previous findings from the SL task in the English native group, the different population—the Korean native speaker group—was also sensitive to the statistical cues necessary for comprehending language.

5.2.2 Korean self-paced reading task: Processing Korean topicalized sentences

For the self-paced reading task, mean comprehension accuracy rates were calculated from the percent of correct responses. The mean comprehension accuracy rate was 86% ($SD = 13$)—87% ($SD = 13$) for the topicalized subject condition, and 85% ($SD = 17$) for the topicalized object condition. The paired $t$-test showed no significant difference between the mean comprehension accuracy rates of the topicalized subject and topicalized object conditions [$t(113) = -1.021, p = .310$], indicating no contrast in comprehension accuracy between topicalized subject and object sentences.
Figure 5.1 describes the profile of residual reading times (RTs) for Korean topicalized subject and object sentences. Overall, Korean native speakers read the topicalized object sentences significantly slower than the topicalized subject sentences across all regions, except for region 5 and 7 \(t(113) = -1.672, p = .097, t(113) = -.306, p = .760,\) respectively]. This suggest that topicalized object sentences impose a greater processing burden on readers compared to the topicalized subject sentences, presumably due to a greater distance between the fronted object and its original position.

Let us look at residual RTs of all regions in detail. At region 1, there is a significant difference between topicalized subject and object sentences, showing slower RTs in the topicalized object condition; the structural case of the fronted object (i.e., -(I)ul, an accusative marker) is likely to create an additional processing cost \(t(113) = 2.411, p = .018\). However, because NPs with different case markers were compared, differences in RTs at region 1 could be due to potentially intrinsic differences between a topic and an accusative case marker rather than to the scrambling (see Miyamoto & Takahashi, 2002).
At regions 2 and 3, there is a significant difference between topicalized subject and object sentences, with slower RTs in the topicalized pattern \([t(113) = 4.023, p = .0001, t(113) = 3.809, p = .0001, \) respectively].

To compare the difficulty of English ORs with that of topicalized object sentences, I focused on results from the topicalized object sentences. The primary regions of interest in the topicalized object condition are regions 4, 5, and 6. At region 4 (the word after the subject), the mean residual RT in the topicalized object condition was significantly slower than that in the topicalized subject condition [a paired sample t-test, \(t(113)= 7.377, p = .0001\]. This is consistent with the idea that the processor acts to establish a filler-gap dependency for the fronted object at the first position where such a relationship might exist—i.e., right after the subject. At region 5 (the locative), there is no significant
difference in mean RT between topicalized subject and topicalized object sentences 
\[ t(113) = -1.672, p = .097 \], even though the mean RT in the topicalized object condition is numerically faster than that in the topicalized subject condition. At region 6, the mean RT of the topicalized object sentences is significantly faster than that of topicalized subject 
\[ t(113) = -4.908, p = .001 \]. This may indicate the dependency formed by the scrambling has already been resolved prior to region 5.

Even though there is no contrast between the mean comprehension accuracy rates in topicalized subject and object conditions, the RT profiles show that overall RTs in the topicalized object condition are significantly slower than those in the topicalized subject condition 
\[ t(113) = 4.317, p = .0001 \]. This suggests that participants had more difficulty processing Korean topicalized object sentences, as reported with object relative clauses in English.

5.3.3 Combined results: Correlation between SL and Korean online comprehension tasks

One of the research questions that the present study was designed to address is whether Korean native speakers’ statistical learning ability affects their performance in processing Korean topicalized object sentences. Because the SL task and processing of topicalized object sentences both involve nonadjacent dependencies, the prediction is that good statistical learners will perform better on Korean topicalized object sentences than poor statistical learners, with faster reading times at the critical regions. This section will focus on the comprehension accuracy and residual reading times of topicalized object sentences.
For the following analysis, the Korean native speakers were divided into two groups (good SLers vs. poor SLers) based on the cut-off point of 90 in the SL task. Based on the comprehension data from the self-paced reading task, the mean comprehension accuracy rates were calculated for the topicalized subject and object conditions, but only the data from topicalized object conditions will be considered here. The overall mean comprehension accuracy rate was 86\% (SD = 13), with 84\% (SD = 8) for fillers. Figure 5.2 shows the mean comprehension accuracy rates of good and poor statistical learners. For the topicalized object condition, the mean comprehension accuracy rates of good and poor statistical learners were 89\% (SD = 15) and 83\% (SD = 19), respectively, with no significant difference between the two groups [t(112) = 1.589, \( p = .116 \)]. Based on these results, there appears to be no effect of statistical learning ability on comprehension accuracy in the topicalized object condition. Regardless of the degree of statistical learning ability, good and poor statistical learners did well.

A possible explanation for this result is that case provides information about an NP’s function within the sentence, usually by indicating its thematic role and/or grammatical relation (see O’Grady 2013:93). This in turn could lead to similar comprehension accuracy rates in both groups.
Figure 5.2 Mean comprehension accuracy of Korean topicalized object sentences for Korean native speakers

Figure 5.3 illustrates residual RTs of Korean topicalized sentences read by good and poor statistical learners. Overall, processing of topicalized object sentences took longer than that of topicalized subject sentences, regardless of whether readers were good or poor statistical learners (see Appendix J). Let us look at the regions of interest (i.e., regions 4, 5, and 6). As seen in Figure 5.3, for the topicalized object condition, good and poor statistical learners also showed no significant difference at the critical regions 4 and 6 [an independent sample t-test, $t(112) = -1.280, p = .203$, $t(112) = .388, p = .699$, respectively].
Figure 5.3 Korean native speakers’ residual RTs profile by good and poor SLers in topicalized object sentences
However, at region 5, there is a significant difference in RTs between good and poor statistical learners in the topicalized object condition \[ t(112) = 1.951, p = .05 \]: good statistical learners read faster at region 5 (the position of the spillover effect\(^{13}\)) compared to poor statistical learners. This points to extra integration cost for connecting a previous word (i.e., a fronted object marked with an accusative marker) with its gap in topicalized object patterns. RTs at region 5 show that in the topicalized object condition, good statistical learners found it relatively easy to keep track of the filler-gap dependency.

We expect a significant difference between good and poor statistical learners in the topicalized object condition at region 4, (ecey ‘yesterday’), the first position after the

---

\(^{13}\)There are often spillover effects in reading times where the effect of difficulty is apparent only one or even two words after the locus of difficulty in L2 sentence processing. It is therefore standard to examine the reading times of following words to identify differences in reading time (Juffs, 1998:415).
subject (*Choy kyoswu-nun* ‘Professor Choi’), but no such difference was found, despite the existence of a numerical difference in the right direction between good and poor statistical learners. We did, however, find a significant difference in RTs between good and poor statistical learners at the next position, region 5 (*samusil-eyse* ‘in his office’), with the good statistical learners reading significantly faster. This can reasonably be interpreted as a spillover effect.

### 5.3 Discussion

The results of Experiment 3 show that Korean native speakers have difficulty processing Korean topicalized object sentences compared to topicalized subject sentences, as shown in Figure 5.1. This supports the hypothesis that processing long-distance topicalization imposes a greater processing load and memory cost, which is consistent with findings from previous studies of this phenomenon (Nakayama, 1995; Mazuka, Itoh, & Kondo, 1997; Miyamoto & Takahashi, 2002). However, the difficulty of topicalized object sentences was not reflected in comprehension accuracy rates: both good and poor groups of Korean native speakers did equally well in the topicalized object condition, indicating that structural case – for example, topic and accusative markers – may provide the information necessary to infer the grammatical role of the fronted NP, leading to good comprehension in topicalized object condition.

The results of this experiment also show that nonadjacent statistical learning ability is associated with processing Korean topicalized object sentences. As predicted, good statistical learners demonstrated faster reading, whereas poor statistical learners read more slowly at the region where a long-distance dependency is resolved. These results
seem to support the finding that statistical learning ability affects comprehension accuracy. Let us discuss these findings in order.

First, even though we found no significant difference in comprehension accuracy rate between good and poor statistical learners in the topicalized object condition, there was a significant difference between good and poor statistical learners in RTs, as illustrated in Figure 5.4. Good statistical learners read significantly faster at region 5 when reading topicalized object sentences than poor statistical learners did, indicating that good statistical learners resolve the filler-gap dependency at the spillover position more quickly than poor statistical learners do. A plausible conclusion is that statistical learning ability affects the processing of Korean topicalized object sentences.

Second, I found no correlation between comprehension accuracy and statistical learning ability. As mentioned above, the additional factor of case marking may facilitate comprehension of the topicalized object sentences, allowing both good and poor statistical learners to show comparably good comprehension accuracy rates.

One factor that needs to be addressed in L2ers’ performance on their second language processing is potential L1 influence. Compared to the results from Korean native speakers as summarized in the present chapter, we could not find an L1 influence in Korean-speaking L2 learners’ performance in processing English ORs (see chapter 4). For example, while good statistical learners within the L2 group read significantly slower at the critical region in reading ORs than poor statistical learners did, good statistical learners within the L2 group read significantly faster at the critical region when reading topicalized object sentences than poor statistical learners did.
Overall, the main findings of this experiment with Korean native speakers confirmed the findings from Misyak & Christiansen’s (2012) study by testing the effect of statistical learning ability on a different language and a different population. As we have seen, there are similar results for both languages and both L1 groups. Although I found no effect of statistical learning ability on comprehension accuracy in either the English or Korean self-paced reading tasks, there was a correlation between statistical learning ability and reading times. These findings contribute to the literature on statistical learning by establishing its relevance to a different group as well as a different language.
CHAPTER 6
GENERAL DISCUSSION AND CONCLUSION

In this chapter, I will discuss to what extent the results obtained from the experiments involving the L1 English group, L2 English group, and L1 Korean group offer insights into various linguistic issues. Based on the results of these three experiments, I will first examine whether the current findings help to shed light on the relation between statistical learning (SL) and language acquisition. I will then turn to the implications for second language acquisition (SLA) and L2 sentence processing.

6.1 Statistical learning and language acquisition

As we saw in chapter 2, only a few prior SL studies have argued that statistical learning is relevant for acquiring language at the syntactic level, and those studies have tested only native speakers of English. Given the difficulty of concluding that the learning mechanism used in statistical learning extends to learning a natural language, it is of interest to see whether the experimental findings reported here can offer empirical support for that idea, at least in the case of nonadjacent dependencies.

Two important findings from the three experiments presented in previous chapters contribute to the study of statistical learning and language acquisition: (1) language learners are capable of extracting distributional cues from input relevant to the existence of non-adjacent dependencies, and (2) this capacity is correlated with learners’ comprehension of patterns involving long-distance dependencies involving relative clauses. Let us consider each finding in turn.
First, results from the three experiments show that both English native speakers and Korean native speakers are capable of extracting simple long-distance dependencies from sequential input by attending to distributional patterns. These results are consistent with findings of previous studies (Gómez, 2002; Misyake et al., 2010; Misyak & Christiansen, 2012). With regard to this study’s nonadjacent SL task, Gómez (2002) notes that in the high-variability condition, two-thirds of her participants showed perfect accuracy (≥ 95%) on the grammaticality test portion of the SL task. Misyak & Christiansen (2012) also reported that one-third of their statistical learners achieved 90% accuracy. In a similar SL task involving a nonadjacent dependency, I also observed that one-third of the statistical learners had accuracy rates on the grammaticality test of 90% or more across two groups – English native speakers and Korean native speakers. Moreover, I observed similar overall accuracy rates on grammaticality judgments in the SL task: 69.2% accuracy rates for the nonadjacent SL task in Misyak & Christiansen’s (2012) study compared to 64% for the English native group and 68% for the Korean native group in my study. Based on the findings from the three experiments reported here, it can be concluded that not only English native speakers but also Korean native speakers are sensitive to statistical cues.

Second, as an initial investigation into a possible relationship between nonadjacent statistical learning and the ability of L2 learners to deal with long-distance dependencies in English, the results indicate an apparent correlation. As reported in chapter 3, Experiment 1 examined the effect of SL ability (measured through learning an artificial grammar) by investigating whether English native speakers are sensitive to distributional cues from sequential input. It was expected that good and poor statistical learners would
have different accuracy rates on comprehension questions and different reading times when resolving a long-distance dependency in the processing of object relative clauses (ORs). Reactivating the filler in the gap position of an OR is more costly in terms of memory load than the parallel reactivation in a subject relative clause (SR) because the filler-gap distance is greater in the former than the latter; this difference in integration cost may cause good statistical learners to differ from poor statistical learners in comprehension accuracy and reading speed. The results of Experiment 1 replicate previous findings that greater success on the SL task correlates with a decrease in reading speed when comprehending ORs; in other words, good statistical learners read faster than poor statistical learners at the region of dependency resolution in sentences containing ORs (Misyak & Christiansen, 2009).

In order to confirm and extend the findings of Experiment 1, a different population and a different language were tested in Experiments 2 and 3. Experiment 2 tested Korean learners of English using the same design as that used in Experiment 1, whereas Experiment 3 tested Korean native speakers’ SL ability and their processing of sentences in their native language. The major findings from these three experiments indicate that reliable effects of SL ability are observed in sentence processing in both languages: in both the English and Korean groups, participants with higher grammaticality accuracy rates on the SL task read sentences that included long-distance dependencies significantly faster than those with lower grammaticality accuracy rates did. The finding of an effect of SL ability in a language other than English helps confirm the existence of a relationship between statistical learning and natural language processing.
The most interesting finding from this dissertation is that statistical learning ability is relevant to learners’ language comprehension. Although the literature on statistical learning has long assumed that SL ability is associated with acquiring natural language, few studies have tested this. The recent work by Misyak and her colleagues directly tested this assumption using English native speakers. Their findings demonstrate a significant correlation between SL ability and L1 sentence processing, but they did not test this correlation in a second language. The present study helps to fill this research gap by finding a relationship between SL ability and sentence processing in L2 learning, showing that good statistical learners outperformed poor statistical learners in at least one type of nonadjacent dependency. Therefore, it is possible to conclude that an effect of SL ability has been observed in a different population and in a different language. Let us now turn to implications for L2 acquisition and L2 processing.

6.2. Second language processing

Sentence processing is a cognitively demanding task. The processor must process words rapidly and linearly and, at the same time, must consider different linguistic information to accurately comprehend sentences. Processors seem to be constrained by factors such as the working memory system.

Based on these observations, memory constraints have played an important role in the study of the processing of syntactically complex sentences (e.g., sentences involving long-distance dependencies). Related research has mainly considered working memory as a factor for explaining individual differences in language comprehension. Despite the
discovery of reliable effects of working memory on L1 sentence processing, similar effects on L2 sentence processing have been questioned in the literature.

Matters may be more promising in the case of SL ability, as measured by the ability to abstract a rule from an invented language. At least in the case of processing nonadjacent dependencies, two aspects of working memory resources are necessary: (i) storage of the structure being built, and (b) integration of incoming and previous words. For example, learners must use their memory resources to extract a specific dependency pair (i.e., pel – jic in [1]) while completing an SL task. Learners need to hold the first word (*pel*) in memory while processing the final word (*jic*), and while connecting incoming words with the previous words so as to establish the correct dependency between *pel* and *jic*.

(1) **pel rop gik jic**

In this way, research on statistical learning might play an important role in understanding the nature of the sentence processing mechanism. Of course, individuals have different memory resource capacities, and this may result in different levels of achievement in learning language. The core questions of this dissertation concern are concerned with why individual L2 learners achieve different levels of ultimate attainment and how L2 learners process sentences. On the descriptive level, the findings reported here confirm those of previous studies that show both English native speakers and L2 learners of English have difficulty processing English ORs.
Studies of L1 sentence processing have provided a possible explanation for this difficulty by showing an association with processing and memory cost. For example, King and Just (1991) examined the relation between reading span and processing ORs by English native speakers and found that processing ORs requires greater working memory than processing SRs: English native speakers with high reading span read significantly faster when reading ORs than did those with low reading span. Gibson and his colleague (Gibson, 1998, 2000; Grodner & Gibson, 1995) also suggest that the greater distance between the head noun and its original position in an OR makes ORs more difficult to process than SRs. According to Gibson’s Dependency Locality Theory, as the number of intervening discourse referents between objects forming a dependency increases, so does the demand on memory resources and the interference with building the necessary connection.

Consistent with the existence of a relationship between L1 processing and memory cost, results from Experiment 1 show (i) that reading times slowed down at the critical region when reading ORs, reflecting the greater linear distance between the extracted direct object and its original position, and (ii) that English native speakers with high grammaticality accuracy rates on the SL task read ORs significantly faster than did those with low grammaticality accuracy rates. One prediction for Experiment 1 that was not borne out, however, was that English native speakers with high grammaticality accuracy rates on the SL task would comprehend ORs significantly better than English native speakers with low grammaticality accuracy rates—in fact, both groups did equally well. Nonetheless, the finding of a relation between grammaticality accuracy rates on the SL task and reading times in the OR condition implies that SL ability is associated with
processing unambiguous syntactic structures, at least in the case of long-distance dependencies.

(a) King and Just (1991)’s study  
(b) Misyak and Christiansen (2010)’s study

![Graph showing mean reading times for subject and object relative clauses](image1)

**Figure 6.1 Mean RTs of subject and object relative clauses from related studies**

Moreover, this finding, as depicted in Figure 6.2, is consistent with those of previous studies, shown in Figure 6.1, in which a significant difference in reading times was observed in ORs but not in SRs. Similarly, results from Experiment 1 show a significant difference in reading times at the embedded verb of ORs. Thus, it is possible to conclude that both SL ability and working memory are closely associated with performance in comprehending language.

---

14 Reading times shown in Figure 6.1 were based on the raw reading times, while the length-adjusted reading times (i.e., residual reading times) in Figure 6.2 were used. With respect with this, the difference of reading times between two Figures does not matter.
To summarize the findings from Experiments 2 and 3, these two experiments showed (i) that like English native speakers in Experiment 1, L2 learners also had difficulty processing ORs, (ii) that L2 learners who performed well on the SL task achieved significantly better comprehension of ORs than did L2 learners who performed poorly on the SL task; (iii) regardless of the L2 learners’ level of proficiency – advanced or intermediate – variation on the SL task was related to different comprehension accuracy on ORs; and (iv) that no influence of L2ers’ L1 on sentence processing was observed. The major finding of this dissertation is that an effect of SL ability was systematically observed for L2 sentence processing. This is particularly interesting because past research has not found a consistent effect of working memory on L2 sentence processing.

However, two puzzles were raised by the experiments reported in this dissertation. First, L2ers’ performance in processing ORs was better than that of English native speakers (91% vs. 73%). As shown in Table 6.1, both groups did well on SR comprehension; however, among good statistical learners, L2 learners achieved higher
comprehension accuracy in reading ORs than did the English L1 group. This question remains unanswered and is therefore a possible avenue for future research.

Table 6.1 Descriptive statistics for L1 and L2 groups

<table>
<thead>
<tr>
<th></th>
<th>L2 group</th>
<th>L1 group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
<td>OR</td>
</tr>
<tr>
<td><strong>Comprehension accuracy (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good SLers</td>
<td>86%</td>
<td>91%</td>
</tr>
<tr>
<td>Poor SLers</td>
<td>82%</td>
<td>75%</td>
</tr>
<tr>
<td><strong>Mean reading times (ms)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good SLers</td>
<td>15.05</td>
<td>147.83</td>
</tr>
<tr>
<td>Poor SLers</td>
<td>50.54</td>
<td>98.08</td>
</tr>
<tr>
<td>Region 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good SLers</td>
<td>121.37</td>
<td>203.03</td>
</tr>
<tr>
<td>Poor SLers</td>
<td>96.35</td>
<td>144.28</td>
</tr>
<tr>
<td>Region 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good SLers</td>
<td>51.92</td>
<td>81.73</td>
</tr>
<tr>
<td>Poor SLers</td>
<td>35.82</td>
<td>26.76</td>
</tr>
</tbody>
</table>

*Note.* Good SLers – good statistical learners; Poor SLers – poor statistical learners

A second puzzle concerns the finding that good Korean-speaking statistical learners read English ORs more slowly than poor Korean-speaking statistical learners across three critical regions, as seen in Table 6.1, whereas good English-speaking statistical learners read ORs faster than poor English-speaking statistical learners at regions 6 and 7. As predicted, among the L1 group, good statistical learners read ORs significantly faster at region 6 than poor statistical learners did, but among the L2 group – particularly among good statistical learners – significantly slower reading times at region 8 were observed. That is, the results from Experiment 2 show that slower reading times correspond to higher comprehension accuracy. As shown in Table 6.1, slower reading time at region 8 in ORs (81.73 ms vs. 27.76 ms) correlates with better comprehension accuracy of ORs (91% vs. 75%), and faster reading time correlates with lower comprehension. Further
research is necessary to investigate why good statistical learners read more slowly when reading ORs.

In this dissertation, I have mainly investigated the effect of statistical learning on L2 learners’ processing of ORs. The effect of SL ability generally appeared in the processing of ORs. The results show that L2ers who did well in the SL task also did well in comprehending ORs, and L2ers who were poor statistical learners also had lower comprehension accuracy. These findings have various implications for second language acquisition. Although the literature on sentence processing and acquisition has focused on working memory (King and Just, 1991; Just & Carpenter, 1992; MacDonald, Just, & Carpenter, 1992; Juffs, 2004; for a review, see Juffs and Harrington, 2011), no reliable effects of this resource have previously been observed in L2 sentence processing.

The results from both groups tested here suggest that statistical learning can help explain differences in sentence processing ability. This raises the possibility that SL may be the long-sought component of working memory capacity relevant to understanding differences in L2 learning ability. Of course, SL ability is not the only factor for explaining why some L2 learners learn more quickly and more successfully than others, but its role in successful processing and acquisition of L2 learners is worthy of further consideration.

Finally, a word of caution is in order. As pointed out to me by William O’Grady (pers. comm.), the sort of nonadjacent dependency used in the SL learning task resembles the sort of dependency found in relative clauses is just one way: elements intervene between the two terms in the dependency. However, the two types of dependency are different in at least three respects:
1) The dependency in a relative clause involves a filler and a gap—not two overt elements.\textsuperscript{15}

2) The dependency in a relative clause is interrupted by varying types and numbers of elements, each with their own processing demands as in examples (2). In contrast, the dependency in the SL task, involved precisely two elements in each training item and test item.

\[(2) \quad \text{a. The man [that Mary met__]}
\]
\[\quad \text{b. The man [that Mary often met__]}
\]
\[\quad \text{c. The pencil [that a friend of Mary’s bought__]}
\]
\[\quad \text{d. The pencil [that I think [Mary bought__]]}
\]

3) The dependencies in natural language are of interest and relevance only to the extent that they contribute to the creation and interpretation of form-meaning mappings. There are no such mappings in the case of the SL task, in which the test items and training items have no meaning.

This does not mean that the correlations discovered here are spurious. To the contrary, it implies that even more robust and telling correlations may be uncovered as improvements in the design of SL tasks create processing demands more similar to those associated with the acquisition and use of natural language.

\textsuperscript{15} In this regard, the pattern used in the SL task is arguably more like the dependency found in languages that use resumptive pronouns in the relative clauses (e.g., the equivalent of The man who John saw him, which is acceptable in languages such as Hebrew). Even there, though, there is an important difference as resumptive pronouns typically match the head noun in particular features (e.g., person, number, and gender), for which there is no counterpart in the SL task.
6.3 Concluding remarks

The research reported in this dissertation investigated the correlation between SL ability and sentence processing that include long-distance dependencies by testing English native speakers, Korean learners of English, and Korean native speakers. Generally, it was found that SL ability is related, to some extent, to aspects of L1 English online processing, L2 English online processing, and L1 Korean online processing. However, further studies with different types of linguistic phenomena (e.g., island effects, local and non-local agreement) are necessary to investigate more thoroughly the suggestion made in this research.
APPENDIX A. ARTIFICIAL LANGUAGE OF SL TASK

72 strings (24 string × 3 dependency-pairs)

1 pel rop gik jic
2 vot poy juf tood
3 dak vev fuf rud
4 pel tib lut jic
5 vot wat dap tood
6 dak siz bul rud
7 pel suv nom jic
8 vot dyt cag tood
9 dak seb kif rud
10 pel bup lig jic
11 vot gos tol tood
12 dak ven lod rud
13 pel bov gyb jic
14 vot deg kem tood
15 dak kav mup rud
16 pel fok kes jic
17 vot doz ned tood
18 dak zaf kuc rud
19 pel gol nim jic
20 vot lat lec tood
21 dak gug fud rud
22 pel pag gaz jic
23 vot lyz gof tood
24 dak fet pez rud
25 vot rop gik tood
26 dak poy juf rud
27 pel vev fuf jic
28 vot tib lut tood
29 dak wat dap rud
30 pel siz bul jic
31 vot suv nom tood
32 dak dyt cag rud
33 pel seb kif jic
34 vot bup lig tood
35 dak gos tol rud
36 pel ven lod jic
37 vot bov gyb tood
38 dak deg kem rud
39 pel kav mup jic
40 vot fok kes tood
41 vot doz ned rud
42 dak doz ned rud
43 pel zaf kuc jic

110
<table>
<thead>
<tr>
<th></th>
<th>Grammatical (novel)</th>
<th>pel</th>
<th>bap</th>
<th>yik</th>
<th>jic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>rop</td>
<td>gik</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Grammatical (novel)</td>
<td>dak</td>
<td>poy</td>
<td>juf</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>vev</td>
<td>fuf</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>tib</td>
<td>lut</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Grammatical (novel)</td>
<td>dak</td>
<td>wat</td>
<td>dap</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>siz</td>
<td>bul</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>suv</td>
<td>nom</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Grammatical (novel)</td>
<td>dak</td>
<td>dyt</td>
<td>cag</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>seb</td>
<td>kif</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>bup</td>
<td>lig</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>bap</td>
<td>yik</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>dev</td>
<td>noz</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Grammatical (novel)</td>
<td>dak</td>
<td>hol</td>
<td>zot</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>vel</td>
<td>fic</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>wal</td>
<td>ros</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Grammatical (novel)</td>
<td>dak</td>
<td>lof</td>
<td>vil</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>maz</td>
<td>tuz</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Grammatical (novel)</td>
<td>vot</td>
<td>moy</td>
<td>bon</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Grammatical (novel)</td>
<td>dak</td>
<td>rus</td>
<td>nex</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Grammatical (novel)</td>
<td>pel</td>
<td>feb</td>
<td>gan</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>ryd</td>
<td>bep</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Ungrammatical</td>
<td>vot</td>
<td>gib</td>
<td>tir</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Ungrammatical</td>
<td>dak</td>
<td>piz</td>
<td>wug</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>reg</td>
<td>lis</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Ungrammatical</td>
<td>vot</td>
<td>wos</td>
<td>zuf</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Ungrammatical</td>
<td>dak</td>
<td>lin</td>
<td>reg</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>tud</td>
<td>jub</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Ungrammatical</td>
<td>vot</td>
<td>jou</td>
<td>var</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Ungrammatical</td>
<td>dak</td>
<td>nen</td>
<td>vuv</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>lyl</td>
<td>yaz</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Ungrammatical</td>
<td>vot</td>
<td>gos</td>
<td>tol</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Ungrammatical</td>
<td>dak</td>
<td>ven</td>
<td>lod</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>bov</td>
<td>gyb</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Ungrammatical</td>
<td>vot</td>
<td>deg</td>
<td>kem</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Ungrammatical</td>
<td>dak</td>
<td>kav</td>
<td>mup</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>fok</td>
<td>kes</td>
<td>tood</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Ungrammatical</td>
<td>vot</td>
<td>doz</td>
<td>ned</td>
<td>rud</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Ungrammatical</td>
<td>dak</td>
<td>zaf</td>
<td>kuc</td>
<td>jic</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Ungrammatical</td>
<td>pel</td>
<td>g0l</td>
<td>nim</td>
<td>tood</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C. TEST SENTENCES AND COMPREHENSION QUESTIONS
OF EXPERIMENT 1 AND 2

1 SR The professor that criticized the student went on vacation in Texas.
   OR The professor that the student criticized went on vacation in Texas.

2 SR The pilot that photographed the crew asked for a date yesterday.
   OR The pilot that the crew photographed asked for a date yesterday.

3 SR The policeman that passed the thief hit someone on the road.
   OR The policeman that the thief passed hit someone on the road.

4 SR The doctor that helped the nurse had a blind date yesterday.
   OR The doctor that the nurse helped had a blind date yesterday.

5 SR The secretary that contacted the boss finished the project in time.
   OR The secretary that the boss contacted finished the project in time.

6 SR The director that called the actor sent candies for Valentine’s Day.
   OR The director that the actor called sent candies for Valentine’s Day.

7 SR The waiter that avoided the customer dropped glasses on the floor.
   OR The waiter that the customer avoided dropped glasses on the floor.

8 SR The child that teased the babysitter had delicious cookies at home.
   OR The child that the babysitter teased had delicious cookies at home.

9 SR The novelist that admired the poet wrote two masterpieces last year.
   OR The novelist that the poet admired wrote two masterpieces last year.

10 SR The skater that watched the runner received medals at the Olympics.
    OR The skater that the runner watched received medals at the Olympics.

11 SR The photographer that telephoned the model worked for a national magazine.
    OR The photographer that the model telephoned worked for a national magazine.

12 SR The manager that ignored the assistant spent the weekend at work.
    OR The manager that the assistant ignored spent the weekend at work.

13 SR The prisoner that pushed the guard stole rings from the store.
    OR The prisoner that the guard pushed stole rings from the store.

14 SR The lawyer that welcomed the client visited the courthouse last week.
    OR The lawyer that the client welcomed visited the courthouse last week.

15 SR The florist that recommended the gardener delivered flowers for a wedding.
    OR The florist that the gardener recommended delivered flowers for a wedding.

16 SR The pianist that approached the violinist played a waltz at night.
    OR The pianist that the violinist approached played a waltz at night.
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SR</td>
<td>Did the professor go on vacation in Texas?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the student criticize the professor?</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>SR</td>
<td>Did the crew photograph the pilot?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the crew ask for a date?</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>SR</td>
<td>Did the policeman hit someone?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the thief pass the policeman?</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>SR</td>
<td>Did the nurse help the doctor?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the nurse have a blind date?</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>Did the secretary finish the project?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the boss contact the secretary?</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>Did the actor call the director?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the actor send candies?</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>SR</td>
<td>Did the waiter drop glasses?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the customer avoid the waiter?</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>SR</td>
<td>Did the babysitter tease the child?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the babysitter have delicious cookies?</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>SR</td>
<td>Did the novelist write two masterpieces?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the poet admire the novelist?</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>SR</td>
<td>Did the runner watch the skater?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the runner receive medals at the Olympics?</td>
<td>No</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>Did the photographer work for a national magazine?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the model telephone the photographer?</td>
<td>Yes</td>
</tr>
<tr>
<td>12</td>
<td>SR</td>
<td>Did the assistant ignore the manager?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the assistant spend the weekend at work?</td>
<td>No</td>
</tr>
<tr>
<td>13</td>
<td>SR</td>
<td>Did the prisoner steal rings?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the guard push the prisoner?</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>SR</td>
<td>Did the client welcome the lawyer?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the client visit the courthouse last week?</td>
<td>No</td>
</tr>
<tr>
<td>15</td>
<td>SR</td>
<td>Did the florist deliver flowers?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the gardener recommend the florist?</td>
<td>Yes</td>
</tr>
<tr>
<td>16</td>
<td>SR</td>
<td>Did the violinist approach the pianist?</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>OR</td>
<td>Did the violinist play a waltz?</td>
<td>No</td>
</tr>
</tbody>
</table>
APPENDIX D. FILLER SENTENCES AND COMPREHENSION QUESTIONS OF EXPERIMENT 1 AND 2

1. Steven called Jennifer and invited her to be in his film.
2. Todd questioned the teacher and left the classroom for another class.
3. Jane had been working for the CIA, but decided to quit.
4. Lisa bought a blue bag, and Mary bought a red one.
5. James did his work and prepared for a trip to Europe.
6. Alex was not only handsome, but he was also quite brilliant.
7. Both the soccer team and the swimming team were doing well.
8. John washed his hands in the restroom and ate the sushi.
9. Harry cleared the Tables in the dining room and left quickly.
10. Erica borrowed the videos from the library, but didn’t return them.
11. The neighbor swept Kelly’s room, but the room was still dirty.
12. The child rode on the back of a camel but cried a lot.
13. The actress received the gifts from the fan and opened them.
14. The singer sang softly and danced a salsa for her fans.
15. Louise wrapped gifts for her cousins and sent them by airmail.
16. Joseph cut out the coupons in the newspaper, but lost them.
17. Although Jenny set her alarm for eight, she slept till noon.
18. The realtor arranged for a meeting although the negotiations were challenging.
19. Sandra moved a cabinet because the office was crowded with chairs.
20. Cathy lifted the box in the living room, but dropped it.
21. Even though Tom drove his car slowly, he got a speeding ticket.
22. Because roses were on sale, Richard sent flowers to his wife.
23. Isaac took care of his sister because his mom was away.
24. According to the student, the professor wrote an excellent recommendation letter.
25. During the long, beautiful sunset, John drew flowers in the park.
26. David bought comics when he got his allowance from his mom.
27. Jim climbed a pyramid in southern Egypt during the summer vacation.
28. Because there was a blackout, Cindy lit the candles last night.
29. Although the puzzle was very difficult, Mrs. Johnson solved it easily.
30. Annie wore the boots last winter because it was so cold.
31. John has decided to resign next year because of personal reasons.
32. Andrea answered every phone call in the office during the boss’s absence.
1. Did Steven call Jennifer? Yes
2. Did the teacher leave the classroom? No
3. Did Jane work for the CIA? Yes
4. Did Mary buy a blue bag? No
5. Did James finish his work? Yes
6. Was Alex quite stupid? No
7. Was the soccer team doing well? Yes
8. Did John eat the pizza? No
9. Did Harry clear the tables? Yes
10. Did Erica return the videos? No
11. Did the neighbor sweep Kelly’s room? Yes
12. Did the child smile? No
13. Did the actress receive the gifts? Yes
14. Did the singer dance a waltz? No
15. Did Louise wrap the gifts to her cousins? Yes
16. Did Joseph still have the coupons? No
17. Did she wake up early? Yes
18. Did the realtor cancel the meeting? No
19. Did Sandra move a cabinet? Yes
20. Did Cathy move the box safely? No
21. Did Tom drive his car slowly? Yes
22. Did Richard send jewelry to his wife? No
23. Did Isaac babysit his sister? Yes
24. Did the student write an excellent letter? No
25. Was the sunset beautiful? Yes
26. Did David lose his allowance? No
27. Was Jim in Egypt during the summer? Yes
28. Did Cindy light the candles last year? No
29. Was the puzzle very difficult? Yes
30. Was it warm last winter? No
31. Did John decide to resign next year? Yes
32. Was the boss absent in the office? No
### APPENDIX E. KOREAN TEST SENTENCES OF EXPERIMENT 3

<table>
<thead>
<tr>
<th>Topicalized</th>
<th>Test sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Subject</td>
<td>조교수는 다행히 어제 연구실에서 지도학생을 잘 만났다.</td>
</tr>
<tr>
<td>Object</td>
<td>지도학생을 다행히 조교수는 어제 연구실에서 잘 만났다.</td>
</tr>
<tr>
<td>2 Subject</td>
<td>천수는 불행히도 어제 출판기념회에서 사진기자를 잘 만나지 못했다.</td>
</tr>
<tr>
<td>Object</td>
<td>사진기자를 불행히도 천수는 어제 출판기념회에서 잘 만나지 못했다.</td>
</tr>
<tr>
<td>3 Subject</td>
<td>영애는 아무튼 어제 호텔에서 노처녀 친구를 분명히 구제해주었다.</td>
</tr>
<tr>
<td>Object</td>
<td>노처녀 친구를 아무튼 영애는 어제 호텔에서 분명히 구제해주었다.</td>
</tr>
<tr>
<td>4 Subject</td>
<td>선주는 솔직히 추석에 부엌에서 친정올케를 정말 돕고 싶었다.</td>
</tr>
<tr>
<td>Object</td>
<td>친정올케를 솔직히 선주는 추석에 정말 돕고 싶었다.</td>
</tr>
<tr>
<td>5 Subject</td>
<td>민영은 이제 방학때 학원에서 이웃집 여학생을 잘 도와줄 것이다.</td>
</tr>
<tr>
<td>Object</td>
<td>이웃집 여학생을 이제 민영은 방학때 학원에서 잘 도와줄 것이다.</td>
</tr>
<tr>
<td>6 Subject</td>
<td>병수는 특히 어제 병원에서 수간호사를 심하게 놀렸다.</td>
</tr>
<tr>
<td>Object</td>
<td>수간호사를 특히 병수는 어제 병원에서 심하게 놀렸다.</td>
</tr>
<tr>
<td>7 Subject</td>
<td>경태는 보통 매일 학교에서 동네친구를 자주 때린다.</td>
</tr>
<tr>
<td>Object</td>
<td>동네친구를 보통 경태는 매일 학교에서 자주 때린다.</td>
</tr>
<tr>
<td>8 Subject</td>
<td>기정은 다행히 어제 대사관에서 호주출신 사무관을 잘 만났다.</td>
</tr>
<tr>
<td>Object</td>
<td>호주출신 사무관을 다행히 기정은 어제 대사관에서 잘 만났다.</td>
</tr>
<tr>
<td>9 Subject</td>
<td>영준은 불행히도 어제 감옥에서 조직폭력배를 우연히 만났다.</td>
</tr>
<tr>
<td>Object</td>
<td>조직폭력배를 불행히도 영준은 어제 감옥에서 우연히 만났다.</td>
</tr>
<tr>
<td>10 Subject</td>
<td>경아는 아무튼 내일 신문사에서 신입정치부 기자를 잘 도와주고 싶었다.</td>
</tr>
<tr>
<td>Object</td>
<td>신입정치부 기자를 아무튼 경아는 내일 신문사에서 잘 도와주고 싶었다.</td>
</tr>
<tr>
<td>11 Subject</td>
<td>영애는 특이 어제 도서관에서 남자친구를 우연히 만나고 싶었다.</td>
</tr>
<tr>
<td>Object</td>
<td>남자친구를 특이 영애는 어제 도서관에서 우연히 만나고 싶었다.</td>
</tr>
<tr>
<td>12 Subject</td>
<td>연애는 흔히 어제 회사에서 성가대 지휘자를 우연히 만날 수도 있다.</td>
</tr>
<tr>
<td>Object</td>
<td>성가대 지휘자를 흔히 연애는 어제 회사에서 우연히 만날 수도 있다.</td>
</tr>
<tr>
<td>13 Subject</td>
<td>정민은 도네 매일 학교에서 남자친구를 자주 만난다.</td>
</tr>
<tr>
<td>Object</td>
<td>남자친구를 도네 정민은 매일 학교에서 자주 만난다.</td>
</tr>
</tbody>
</table>
APPENDIX F. KOREAN FILLER SENTENCES OF EXPERIMENT 3

1 엄마와 함께 아기돼지들이 다른 모양의 초콜렛을 어제 만들었다.
2 비오는 날 대부분의 너구리가 삐끗한 감자를 할머니와 함께 먹었다.
3 어떤 집에 여러마리의 고양이가 복수하려 호황이 집에 올것이다.
4 아기곰 세마리가 집에서 친구들과 함께 모든 침대에 누워보았다.
5 계곡에서 모든 생쥐가 한개씩 뽑을래를 깨끗히 했다.
6 담장밑에서 모든 동물들이 사탕을 한개씩 먹지 않았다.
7 장기자랑에서 양봉곡으로 선생님들이 한곡 노래만 외치지 않았다.
8 경기에서 열개의 로봇들이 같은 트랙을 다 돌지 않았다.
9 경기에서 열개의 로봇들이 같은 트랙을 다 돌지 않았다.
10 학교에서 세마리의 여우가 같은 문제를 증명하지 않았다.
11 어제 나무 아래서 곰들이가 두가지 소원만 적지 않았다.
12 동짓날 할머니가 여섯친구에게 팥죽을 한그릇씩 주지 않았다.
13 늦은 밤에 강아지 세마리가 고양이 한마리씩 따라가지 않았다.
14 담장밑에서 친구들과 함께 공들이 사탕 한개씩을 모두 골랐다.
15 해질무렵 고양이들이 농장에서 모든 조롱박을 새지 않았다.
16 오늘 유치원에서 토끼마다 선물 두개씩을 상으로 받았다.
17 장난감 가게에서 아기공들마다 다른 종류의 장난감을 다 샀다.
18 숲에서 엄마들과 함께 아기 담장과 두마리는 도토리만 줌지 않았다.
19 어린이날 때지 상형제가 다 행버기를 사지 않았다.
20 문방구에서 모든 여우들이 축구공 두개를 사지 않았다.
21 고개에서 호랑이 향나리는 아기고 싶은 맛에 사내와 싸움을 했다.
22 냉장고에서 강아지 세마리가 열지 한개만 깨먹지 않았다.
23 이른 아침 원숭이 한마리는 놀이공원에서 목마를 타지 않았다.
24 민수가 친구 동준께 감사소림에 대해서 수업에 늦으셨다고 말했다.
25 정수가 친구 준수에게 감사소림에 대해서 연구실로 가셨어요라고 말했다.
26 준우가 친구 준기에 이 선풍기에 대해서 아주 무서우리라고 말했다.
27 수미가 친구 준우에게 자신의 할아버지에 대해서 공원에서 운동하셔라고 말했다.
28 희주가 남동생 우영에게 자신의 할머니에 대해서 병원에 가셨어라고 말했다.
29 선주가 여동생 선혜에게 자신의 아버지에 대해서 회사에 가셨다고 말했다.
30 희진이가 친구 정희에게 박사장님에 대해서 아주 친절하셔라고 말했다.
31 선희가 남동생 상문에게 자신의 큰아버지에 대해서 많이 아프서리라고 말했다.
APPENDIX G. LINGUISTIC BACKGROUND QUESTIONNAIRE
[ENGLISH NATIVE SPEAKERS]

A. General information
1. Participant: __________
2. Gender: [ ] F [ ] M
3. Age: __________
4. Major: __________

B. Language background
1. Native language: __________
2. Mother’s dominant language: __________
3. Father’s dominant language: __________
4. Language(s) spoken at home as a child: __________
5. Language(s) you spoke during the first five years of your life:
   __________
6. Languages you use:
   • At home: _________________
   • At school: _________________
   • At work: _________________
7. What language do you feel most comfortable with at this time?
   __________
APPENDIX H. BACKGROUND QUESTIONNAIRE
[KOREAN LEARNERS OF ENGLISH]

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)
Home Country: ________________
Age: ______________
Major: ______________________
Email address: _______________________

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. Approximately how many hours a day do you use English? (Please specify) ________ hours
6. 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)
7. 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)

8. If you have any official score of English proficiency test, please report them.
   TOEIC ______________________  When: ___________________________
   TOEFL______________________  When: ___________________________
   TEP_________________________  When: ___________________________
APPENDIX I. 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 complete this within 10 minutes.

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.
Answers of C-test

Text 1:
We all live with other people’s expectations of us. These are a reflection of them trying to understand us; they are predictions of what we will think, do and feel. Generally we accept the status quo, but these expectations can be hard to handle when they come from our family and can be difficult to ignore, especially when they come from our parents/partner.

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 obesity has become a serious problem in this country as children consume more sugar-based food and spend less time getting the necessary exercise. Many parents have questioned schools’ decisions to allow vending machines which dispense candy and soft drinks. Many schools, though, have come to rely on the money these machines generate through agreements with the companies which makes soft drinks and junk food.
APPENDIX J. RAW READING TIMES OF EXPERIMENT 1, 2 AND 3

Experiment 1: English native group's raw reading times profile by good and poor SLers.
Experiment 2: All L2ers’ raw reading times profile by good and poor SLers
Experiment 3: Korean native group’s raw reading times profile by good and poor SLers
APPENDIX K. RAW READING TIMES OF ENGLISH AND KOREAN SELF-PACED READING TASKS FROM THREE GROUPS

Experiment 1: English native group’s raw RTs profile for English RCs

a. The novelist that admired the poet wrote two masterpieces last year.
b. The novelist that the poet admired wrote two masterpieces last year.
Experiment 2: Korean L2ers’ raw RTs profile for English RCs

Experiment 3: Korean native group’s raw RTs profile for Korean topicalized sentence
REFERENCES


research. London: Longman.


