PROCESSING AND ACQUISITION OF SCRAMBLED SENTENCES BY LEARNERS OF JAPANESE AS A SECOND LANGUAGE

by

Yasumasa Shigenaga

A Dissertation Submitted to the Faculty of the

GRADUATE INTERDISCIPLINARY DOCTORAL PROGRAM IN SECOND LANGUAGE ACQUISITION AND TEACHING

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2014

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Yasumasa Shigenaga, titled Processing and Acquisition of Scrambled Sentences by Learners of Japanese as a Second Language and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

____ Date: 08/28/2014
Thomas G. Bever

____ Date: 08/28/2014
Kenneth Forster

____ Date: 08/28/2014
Janet Nicol

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

____ Date: 08/28/2014
Dissertation Director: Thomas G. Bever

STATEMENT BY AUTHOR

SIGNED: Yasumasa Shigenaga

ACKNOWLEDGEMENTS

This dissertation could not have been completed without the help and support from many people. First, I would like to thank my advisor, Dr. Thomas Bever. It was his guidance and insights that shaped the ideas about this dissertation. He was always patient with my very slow progress and provided me with constant encouragement, which motivated me to keep on working.

I also would like to thank the other members of my dissertation committee. Dr. Kenneth Forster spared many hours for me especially during the data-analysis phase. Without his support, I could not have analyzed the large amount of data I had collected for this project. Dr. Janet Nicol also spared a lot of her valuable time for me especially when I was struggling with the creation of the experimental stimuli. Her practical advice made the creation of the materials so much easier.

This dissertation was partly funded by the SLAT Research Fund. The SLAT program has been very generous throughout my time at the University of Arizona. I am so glad that I have been part of such a supportive program.

My gratitude is extended to Ryoko Utsunomiya, who drew beautiful pictures for the experiments, and to Ayaka Komori, who audio-recorded the sentence stimuli. I owe thanks to my friends and colleagues in the Department of East Asian Studies, the Department of Linguistics, and the SLAT program for making my time in Tucson so much more enjoyable. Thanks also to John Harris, who was available for moral support and reminded me that life is pretty good.

Lastly, I would like to thank John Bilotta for encouraging me to pursue a higher degree and supporting me throughout my time in the doctoral program.

DEDICATION

To my parents

TABLE OF CONTENTS

LIST OF FIGURES 10

LIST OF TABLES 12

ABSTRACT 14

CHAPTER 1: INTRODUCTION 16
1.1 Introduction 16
1.2 Scrambling in Japanese 17
1.2.1 Structure of Japanese scrambling 18
1.2.2 Use of scrambling 21
1.2.3 Frequency of scrambling 23
1.3 The present study 25
1.3.1 Research questions 26
1.3.2 Overview of the methods 29
1.3.2.1 Participants 29
1.3.2.2 Tasks 33
1.3.2.3 Order of participation 33
1.4 Organization of chapters 34

CHAPTER 2: THE RELATION OF GRAMMATICAL KNOWLEDGE AND PERFORMANCE IN L2 LEARNERS' PRODUCTION OF SCRAMBLED SENTENCES 36
2.1 Introduction 36
2.2 Two concepts that are most relevant to the present study 37
2.2.1 Canonical template 37
2.2.2 Reversibility 39
2.3 Literature review 42
2.3.1 Processing of OSV sentences by L1 Japanese children 42
2.3.2 Processing of OSV sentences by L2 learners of Japanese 49
2.4 Rationale and purposes of the present study 54
2.5 Method for Task 1 57
2.5.1 Participants 57
2.5.2 Materials and design 57
2.5.3 Procedure 60
2.5.4 Scoring and analysis 60
2.6 Results of Task 1 63
2.6.1 Mean response accuracy data based on the “lax” scoring 63
2.6.1.1 Transitive sentences 63
2.6.1.2 Locative sentences 66
2.6.2 Mean response accuracy data based on the “strict” scoring 67
2.6.2.1 Transitive sentences 67
2.6.2.2 Locative sentences 71
2.6.3 Raw frequency data 73
2.6.3.1 Reversible sentences 73

TABLE OF CONTENTS - *Continued*

2.6.3.2 Non-reversible sentences....77
2.6.3.3 Locative sentences....79
2.6.4 Discussion....82
2.7 Method for Task 2....84
2.7.1 Participants....84
2.7.2 Materials and design....84
2.7.3 Procedures....87
2.7.4 Scoring and analysis....88
2.7.4.1 Mean response accuracy data....88
2.7.4.2 Raw frequency data....90
2.7.4.3 Reaction time data....90
2.8 Results of Task 2....91
2.8.1 Mean response accuracy data based on the "lax" scoring....91
2.8.1.1 NNS participants....91
2.8.1.2 NS participants....95
2.8.2 Mean response accuracy data based on the "strict" scoring....96
2.8.2.1 NNS participants....96
2.8.2.2 NS participants....100
2.8.3 Raw frequency data....101
2.8.3.1 NNS participants: Reversible sentences....102
2.8.3.2 NNS participants: Non-reversible sentences....106
2.8.3.3 NS participants....110
2.8.4 Reaction time data....113
2.8.5 Discussion....115
2.9 Grammatical knowledge and production performance....117
2.9.1 Mean response accuracy data....117
2.9.2 Raw frequency data....121
2.9.2.1 Reversible sentences....121
2.9.2.2 Non-reversible sentences....127
2.10 General discussion....130

CHAPTER 3: COMPREHENSION OF SCRAMBLED SENTENCES BY L2 LEARNERS OF JAPANESE....134
3.1 Introduction....134
3.2 The Japanese parser and scrambling....134
3.3 Literature review....137
3.3.1 Psycholinguistic investigations on comprehension of scrambled sentences by NSs of Japanese....137
3.3.2 L2 studies on comprehension of scrambled sentences....143
3.3.3 L2 sentence processing and the Shallow Structure Hypothesis....146
3.4 The present study....149
3.5 Method for the pilot study....151
3.5.1 Participants....151
3.5.2 Materials and design....152

TABLE OF CONTENTS - *Continued*

3.5.3 Procedure 152
3.6 Results of the pilot study 153
3.6.1 Reaction times 153
3.6.2 Accuracy rates 155
3.6.3 Discussion 156
3.7 Method for Task 3 157
3.7.1 Participants 157
3.7.2 Materials and design 157
3.7.3 Procedure 159
3.7.4 Analysis 160
3.8 Results of Task 3 161
3.8.1 The ADV group 163
3.8.2 The HI group 164
3.8.3 The LI group 166
3.8.4 The NS group 168
3.8.5 Discussion 169
3.9 General discussion 170

CHAPTER 4: SYNTACTIC PERSISTENCE IN THE PRODUCTION OF SCRAMBLED SENTENCES BY L2 LEARNERS 174
4.1 Introduction 174
4.2 Literature review 175
4.2.1 Syntactic persistence phenomenon 175
4.2.2 L2 studies of syntactic persistence 180
4.2.2.1 Cross-linguistic studies 181
4.2.2.2 Within-L2 studies 185
4.2.3 Syntactic persistence studies on Japanese scrambling 189
4.3 The present study 190
4.4 Method for Task 4 191
4.4.1 Participants 191
4.4.2 Materials and design 191
4.4.3 Procedures 194
4.4.4 Scoring and analysis 197
4.5 Results of Task 4 198
4.5.1 Discussion 199
4.6 The follow-up tasks (Tasks 5 and 6) 202
4.7 Method for the follow-up tasks 204
4.7.1 Participants 204
4.7.2 Materials and design 205
4.7.2.1 Task 5 205
4.7.2.2 Task 6 206
4.7.3 Procedures 208
4.7.3.1 Task 5 208
4.7.3.2 Task 6 209

TABLE OF CONTENTS - *Continued*

4.7.4 Scoring and analysis 210
4.7.4.1 Task 5 210
4.7.4.2 Task 6 210
4.8 Results of the follow-up tasks 211
4.8.1 Task 5....................211
4.8.2 Task 6....................213
4.8.3 Correct use of particles in Tasks 5 and 6....................214
4.8.4 Results summary of the follow-up tasks 217
4.9 General discussion....................219

CHAPTER 5: SUMMARY AND CONCLUSION....................225
5.1 Summaries of the three articles 225
5.1.1 Chapter 2 225
5.1.2 Chapter 3 227
5.1.3 Chapter 4 228
5.2 Answers to the research questions 229
5.3 Pedagogical implications and future research directions....................234

APPENDIX 1A: GENERAL BACKGROUND INFORMATION OF THE NNS PARTICIPANTS....................237
APPENDIX 1B: NNS PARTICIPANTS' SELF-REPORTED PROFICIENCY AND THEIR SPOT SCORES 239
APPENDIX 2A: ITEMS FOR TASK 1 241
APPENDIX 2B: PICTURES USED FOR TASKS 2, 4, AND 6 243
APPENDIX 3A: ITEMS FOR TASK 3 249
APPENDIX 3B: MEAN RESIDUAL READING TIMES IN TASK 3 AMONG THE 10 PARTICIPANTS WHO SCORED HIGHEST IN SPOT....................252
APPENDIX 4A: PRIME SENTENCES FOR TASK 4 253
APPENDIX 4B: PROPORTIONS (IN PERCENT) OF THE CANONICAL, SCRAMBLED, AND OTHER RESPONSES IN TASK 4 254
APPENDIX 4C: THE SIX ADDITIONAL TARGET PICTURES USED IN TASK 6....................255
APPENDIX 4D: PROPORTIONS (IN PERCENT) OF THE CANONICAL, SCRAMBLED, AND OTHER RESPONSES IN TASK 6 256

REFERENCES 257

LIST OF FIGURES

Figure 2.1: Mean accuracy rates for the transitive sentences based on the “lax” scoring in Task1 63

Figure 2.2: Mean accuracy rates for the locative sentences based on the “lax” scoring in Task 1 66

Figure 2.3: Mean accuracy rates for the transitive sentences based on the “strict” scoring in Task 1 68

Figure 2.4: Mean accuracy rates for the locative sentences based on the “strict” scoring in Task 1 72

Figure 2.5: The number of participants by proficiency group who completed the reversible scrambled sentences correctly in Task 1 75

Figure 2.6: The number of participants by proficiency group who completed the non-reversible scrambled sentences correctly in Task 1 78

Figure 2.7: The number of participants by proficiency group who completed the locative scrambled sentences correctly in Task 1 81

Figure 2.8: Examples of the pictures used in Task 2: reversible canonical/scrambled (above) and non-reversible canonical/scrambled (below) 86

Figure 2.9: Mean accuracy rates by the NNS groups based on the “lax” scoring in Task 2 91

Figure 2.10: Mean accuracy rates by the NS group based on the “lax” scoring in Task 2 95

Figure 2.11: Mean accuracy rates by the NNS groups based on the “strict” scoring in Task 2 97

Figure 2.12: Mean accuracy rates by the NS group based on the “strict” scoring in Task 2 101

Figure 2.13: The number of participants by proficiency group who produced the reversible scrambled sentences correctly in Task 2 104

Figure 2.14: The number of participants by proficiency group who produced the non-reversible scrambled sentences correctly in Task 2 108

Figure 2.15: Reaction times for each sentence type by participant group in Task 2 113

Figure 2.16: Mean accuracy rates of Tasks 1 and 2 for the ADV group 118

Figure 2.17: Mean accuracy rates of Tasks 1 and 2 for the HI group 119

LIST OF FIGURES - *Continued*

Figure 2.18: Mean accuracy rates of Tasks 1 and 2 for the LI group 120

Figure 2.19: The number of participants by group who made a certain numbers of erroneous SOV-based responses in the reversible scrambled condition in Tasks 1 and 2 123

Figure 2.20: The number of participants by group who made a certain numbers of erroneous NP1-subject responses in the reversible scrambled condition in Tasks 1 and 2 124

Figure 2.21: The number of participants by group who made a certain numbers of erroneous NP2-object responses in the reversible scrambled condition in Tasks 1 and 2 124

Figure 3.1: Mean reaction times for each sentence type by the NNS and NS groups (the pilot study) 154

Figure 3.2: Mean accuracy rates for each sentence type by the NNS and NS groups (the pilot study) 155

Figure 3.3: Mean residual reading times for the ADV group in Task 3 163

Figure 3.4: Mean residual reading times for the HI group in Task 3 165

Figure 3.5: Mean residual reading times for the LI group in Task 3 166

Figure 3.6: Mean residual reading times for the NS group in Task 3 168

Figure 4.1: An example of a reversible picture (left) and a non-reversible picture (right) used in Task 4 193

Figure 4.2: Visual presentation of a trial in Task 4 196

Figure 4.3: Utterance rates of scrambled sentences in Task 4 198

Figure 4.4: Mean accuracy rates in Task 5 212

Figure 4.5: Utterance rates of scrambled sentences in Task 6 213

Figure 4.6: Mean accuracy rates for the reversible canonical sentences in Tasks 5 and 6 215

Figure 4.7: Mean accuracy rates for each of the additional 6 items in Tasks 5 and 6 216

Figure 4.8: Mean accuracy rates for the additional 6 items in Tasks 5 and 6 217

LIST OF TABLES

Table 1.1: Summary of NNS participants' background 32

Table 1.2: Ranges and means of the SPOT scores in percent by proficiency group 32

Table 1.3: Type, purpose, and participants of the seven tasks in the present study 33

Table 2.1: Raw frequencies of the particles used for the reversible sentences in Task 1 74

Table 2.2: Raw frequencies of the particles used for the non-reversible sentences in Task 1.......77

Table 2.3: Raw frequencies of the particles used for the locative sentences in Task 1 80

Table 2.4: Raw frequencies of the particles that the NNS participants used to produce reversible scrambled sentences in Task 2...102

Table 2.5: Raw frequencies of the particles that the NNS participants used to produce non-reversible scrambled sentences in Task 2 107

Table 2.6: Raw frequencies of the particles that the NS participants used to produce reversible scrambled sentences in Task 2...111

Table 2.7: Raw frequencies of the particles that the NS participants used to produce non-reversible scrambled sentences in Task 2 112

Table 2.8: The number of participants who provided SOV-based responses and the frequencies of the SOV-based forms in the reversible scrambled condition in Tasks 1 and 2 ...122

Table 2.9: The number of participants who provided SOV-based responses and the frequencies of the SOV-based forms in the reversible scrambled condition in Tasks 1 and 2 (by group) 123

Table 2.10: The list of participants who completed or produced reversible scrambled sentences equal to or less than 50% of the time in Tasks 1 and 2...125

Table 2.11: The number of participants in terms of the frequency of the OSV sentence completion/production in Tasks 1 and 2 (reversible scrambled condition) 125

Table 2.12: Frequencies of the SOV-based form in the non-reversible scrambled condition in Tasks 1 and 2 127

Table 2.13: The list of participants who completed or produced non-reversible scrambled sentences equal to or less than 50% of the time in Tasks 1 and 2............................128

LIST OF TABLES - *Continued*

Table 2.14: The number of participants in terms of the frequency of the OSV sentence completion/production in Tasks 1 and 2 (non-reversible scrambled condition) 128

Table 3.1: Mean residual reading times for each sentence type by participant group in Task 3 162

Table 3.2: The number of trimmed data points for each analysis in Task 3 162

ABSTRACT

The Japanese language exhibits a free word-order phenomenon called scrambling. Because each noun phrase (NP) is case-marked with postpositional particles, it allows a freer word order than such languages as English. For simple transitive sentences, Subject-Object-Verb is the canonical word order while OSV is the scrambled word order. Previous studies with native speaker (NS) children have found that they go through a developmental stage during which they consistently misunderstand scrambled sentences, taking the first NP in OSV sentences to be the subject. It has also been found that NS adults experience slowdowns in reading and comprehending scrambled sentences. However, investigations into the processing of scrambled sentences by second language (L2) learners have been scarce, and it is not entirely clear how scrambled sentences are processed and acquired by L2 learners.

This three-article dissertation aimed at investigating how simple transitive sentences with a scrambled word order (i.e., OSV) are processed and acquired by L2 learners whose native language is English. The first article (Chapter 2) examined L2 learners' grammatical knowledge and production performance of the OSV sentences through two tasks (fill-in-the-blank and picture description). The results indicated a positive relationship between the learners' general proficiency in Japanese and their knowledge/production performance of the OSV sentences, although there was a rather large individual difference even within proficiency groups. It was also found that the difficulty in producing OSV sentences was mostly due to a lack of grammatical knowledge, but the relationship of grammatical knowledge and production performance interacted with the types of sentences. For reversible sentences (in which both the subject and object NPs are animate), there was evidence that errors in the production of OSV sentences were caused by the overuse of the canonical template (i.e., SOV). For non-reversible

sentences (in which the subject NP is animate and the object NP is inanimate), on the other hand, there was little evidence that a processing problem such as the overuse of the SOV template caused the production difficulty.

The second article (Chapter 3) examined the comprehension processes of OSV sentences. While the results of a pilot study (sentence correctness decision task) indicated that both the L2 learners and NSs took longer to read and comprehend OSV sentences than SOV sentences, the results of a self-paced reading task suggested that the processing of OSV sentences by L2 learners might be quite different from that of NSs. The NS participants read more slowly at the second NP position when they read the OSV sentences. On the other hand, the L2 learners, regardless of their proficiency level, did not show such slowdowns. However, the data provided evidence that the advanced L2 learners integrated the case particles more consistently in their sentence comprehension than the learners with lower proficiency.

The third article (Chapter 4) examined whether a psycholinguistic task (syntactic persistence with picture description) might facilitate the production of scrambled sentences among L2 learners, for the purpose of exploring the possibility of using such a method as an L2 instructional tool. While the main task (Task 4, which used regular SOV/OSV sentences as primes) was not very effective in eliciting the production of OSV sentences, the follow-up task (Task 6, which used questions in SOV/OSV orders as primes) observed a more positive effect of syntactic persistence. Based on the results, explicit instruction and practice on scrambling is suggested. Since processing of scrambled sentences requires that L2 learners be aware of the functions of case markers (and other postpositional particles) instead of relying on the canonical template, such explicit instruction and practice may also contribute to the acquisition of the particles that mark case.

CHAPTER 1: INTRODUCTION

1.1 Introduction

Languages differ quite substantially in the way they encode grammatical relations in a clause. English has a relatively strict word order, which provides a reliable clue to the thematic-role assignment. In the sentence "John kicked Mary.", *John* is the agent of kicking and *Mary* is the patient, but not the other way around. In Japanese, on the other hand, the most crucial information on the thematic-role assignment is provided by postpositional particles (case-markers) instead of word order.[1] This feature of Japanese is exemplified in a phenomenon called *scrambling*. While there is a certain canonical word order, the arguments (and adjuncts) can be placed in alternative orders without changing the meaning of the sentence. In the following example of the simple transitive sentences, (1.1a) is the sentence in canonical word order, and (1.1b) is its typical scrambled counterpart:

(1.1) a. John-ga Mary-o ketta.
-Nom -Acc kicked
'John kicked Mary.'

b. Mary-o John-ga ketta.
-Acc -Nom kicked

The noun phrases (NPs), *John* and *Mary*, are marked with the nominative case-marker –*ga* and the accusative case-marker –*o*, respectively, and therefore, the grammatical subject is *John* and the object is *Mary* in both sentences, regardless of word order.

[1] Case-markers (e.g., accusative, dative, nominative) are actually a subcategory of postpositional particles. In the generative framework, case-markers are treated as part of the NP and do not project a separate node (e.g., Miyagawa, 1989). On the other hand, postpositional particles such as *de* ('at') and *kara* ('from') project their own P nodes in the same manner as English prepositions. In the present study, the two terms will be used interchangeably, however.

Given that grammatical relations are encoded primarily by word order in English and by case-markers in Japanese, how do learners of Japanese as a second language (L2) whose first language (L1) is English process scrambling in Japanese? Do they process scrambling in a way similar or comparable to native speakers?

This three-article dissertation was aimed at investigating how simple transitive scrambled sentences in the Object-Subject-Verb (OSV) word order are produced (Chapter 2) and comprehended (Chapter 3) by L1 English learners of Japanese. Additionally, the present study was aimed at examining whether a psycholinguistic task (syntactic persistence during picture description) might facilitate the production of scrambled sentences among L2 learners, for the purpose of exploring the possibility of using such a method as an L2 instructional tool (Chapter 4). Findings in L2 processing research are important not only for the theories of L2 acquisition but also for instructional purposes, as they inform the kinds of processing strategies L2 learners need to acquire to achieve a more efficient and effective processing of an L2. It is hoped that the present study can make a contribution in that direction.

This chapter first provides an overview of the scrambling phenomenon in Japanese (grammatical structure, use, and frequency). It will then present the research questions and the overall methods (participants and order of tasks) of the present study. Finally, the organization of the subsequent chapters is summarized.

1.2 Scrambling in Japanese

As observed above, due to the postpositional marking of NPs, Japanese (among many other languages) exhibits a word-order phenomenon commonly called scrambling. The exception is the verb, which needs to be placed at the end of the clause. Scrambling can be local or long distance. (1.1b) above is an example of local scrambling (commonly referred to as “clause-internal scrambling”) while (1.2b) below is an example of “long distance scrambling”, which exhibits scrambling out of a clause.[2] The example is based on Yamashita (2002: 602).

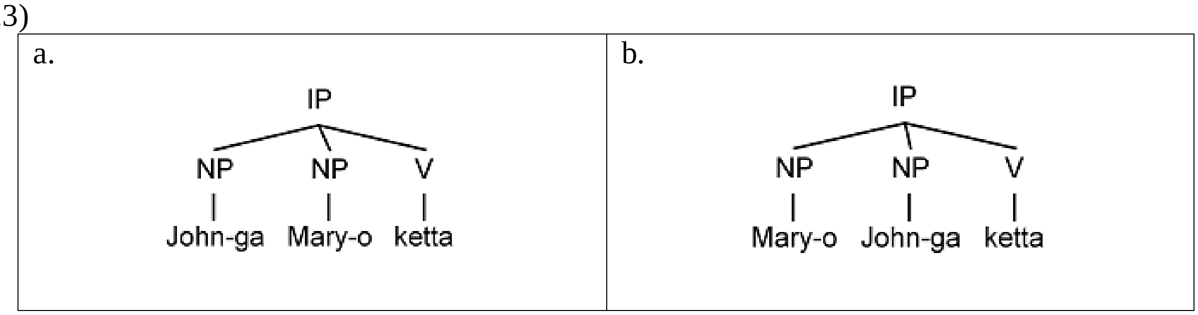
(1.2) a. *Canonical:*
John-wa [Mary-ga ookina ie-o katta-to] omotteiru.
-Top -Nom big house-Acc bought-Comp think
‘John thinks that Mary bought a big house.’

b. *Scrambled:*
Ookina ie-o John-wa [Mary-ga katta-to] omotteiru.
big house-Acc -Top -Nom bought-Comp think

1.2.1 Structure of Japanese scrambling

In some earlier analyses, scrambling was taken as evidence that Japanese has a non-configurational, “flat” structure (e.g., Farmer, 1984; Hale, 1980). In a non-configurational language, a verb phrase (VP) in a clause is not assumed, and therefore, there is no hierarchical difference between the subject and the object. According to such analyses, the canonical sentence (1.1a) and its scrambled counterpart (1.1b) are represented by the reduced tree diagrams in the following manner.

(1.3)

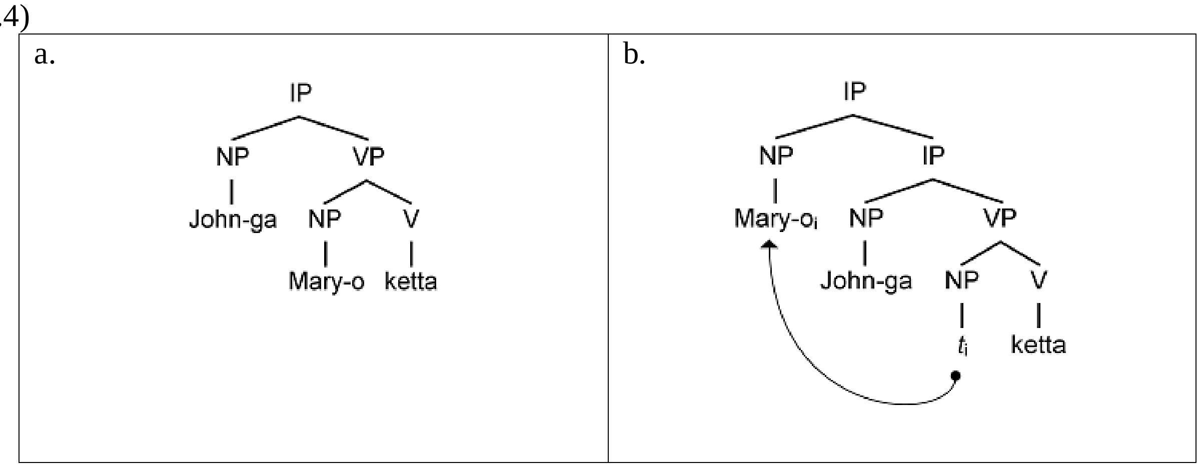


[2] The focus of the present study is the simple transitive *Object – Subject – Verb* (OSV) word order (i.e., the clause-internal type), and hence, the following discussion centers on this particular type of scrambling.

All the arguments are sisters of the verb in (1.3). The apparent advantage of this analysis is that it can easily account for the freer word order of Japanese. However, no word order is assumed to be canonical within this analysis, which goes against native Japanese speakers' intuition that "NP-*ga* NP-*o* V", for example, is the canonical word order (Yamashita, 1997).

More recently, several researchers have suggested that Japanese does have a configurational, hierarchical structure, just as languages, such as English and French, have, and currently it is generally assumed that scrambled sentences are derived from canonical ones by movement. Following this analysis, the canonical sentence (1.1a) and its scrambled counterpart (1.1b) can be represented in the following way using reduced tree diagrams. Note that in (1.4b) an additional IP node is adjoined to the tree, and the accusative NP *Mary* is moved higher in the tree, leaving a gap (trace).

(1.4)



The evidence for the configurational analysis includes such syntactic phenomena as pronominal coreference (Saito, 1985) and weak crossover (Saito & Hoji, 1983) (cf. Nemoto, 1999, for overview). Saito (1985), for instance, maintains that the binding rule, "A pronoun cannot c-command its antecedent" (p. 36), cannot be fulfilled if the existence of a VP node is not assumed. Below are examples used in Saito (1985: 37).

(1.5) a. John$_i$-ga [Mary-ga kare$_i$-ni okutta tegami]-o mada yonde inai (koto)
-Nom -Nom he-Dat sent letter-Acc yet read not
'John has not yet read the letter Mary sent him.'

b. *[Kare$_i$-ga [Mary-ga John$_i$-ni okutta tegami]-o mada yonde inai (koto)
He-Nom -Nom -Dat sent letter-Acc yet read not
'He has not yet read the letter Mary sent John.'

c. [John$_i$-kara okane-o moratta hito]-ga kare$_i$-o suisenshita (koto)
-from money-Acc received person-Nom he-Acc recommended
'The person (who) received money from John recommended him.'

d. [Kare$_i$-kara okane-o moratta hito]-ga John$_i$-o suisenshita (koto)
He-from money-Acc received person-Nom -Acc recommended
'The person (who) received money from him recommended John.'

Saito (1985) argues that, if the existence of a VP is not assumed in Japanese, (1.5c) should be ungrammatical just as is (1.5b). This is so because, without a VP, *kare* ('he') would c-command its antecedent *John* in (1.5c). However, (1.5c) is a well-formed sentence in Japanese. Saito attributes this asymmetry between the subject NP and object NP to the existence of a VP. (1.5d) shows that a pronominal can precede its antecedent in Japanese.

Kuroda (1980) demonstrates that scrambled sentences are derived from canonical ones via movement using the following example. Kuroda observes that the object and the numeral quantifier (NQ) can be separated by the intervening subject, but the subject and the NQ that modifies it cannot be separated by the intervening object. The following example is from Kuroda (1980: 27):

(1.6) a. Utide-no kozuti-o igirisuzin-ga hutatu katta
striking-Gen mallet-Acc Englishman-Nom 2 objects bought
'An Englishman bought 2 *mallets of luck*.'

b. *[Igirisuzin-ga utide-no kozuti-o sannin katta
Englishman-Nom striking-Gen mallet-Acc 3 people bought
'Englishman bought (the) *mallet of luck*, three people.'

(1.6a) is the case for the intervening subject and (1.6b) is the case for the intervening object. While (1.6a) is grammatical, (1.6b) is ill-formed. Based on this contrast, Kuroda maintains that the basic Japanese word order is SOV, and that the scrambled word order derives from the canonical word order via movement.

Saito (1985) also provides evidence for the movement analysis. The examples below are from Saito (1985: 39).

(1.7) a. *[Kare$_i$-ga [Mary-ga John$_i$-ni okutta tegemi]-o mada yonde inai (koto)
he-Nom -Nom -to sent letter-Acc yet read not
'He has not yet read the letter Mary sent to John.'

b. [Mary-ga John$_i$-ni okutta tegami]-o kare$_i$-ga mada yonde inai (koto)
-Nom -to sent letter-Acc he-Nom yet read not
'The letter Mary sent to John, he has not read yet.'

(1.7a), which is identical to (1.5b), is ungrammatical because the pronoun *kare* ('he') c-commands the R-expression *John*. On the other hand, the subject does not c-command the object in (1.7b) and the sentence is well-formed. Thus, it can be taken as evidence that the object moved to a position higher than the subject as illustrated in (1.4b).

1.2.2 Use of scrambling

While it has been argued that the scrambled NP receives some focus or emphasis (e.g., Shibatani, 1990; Yamashita, 2002), what exactly motivates the use of scrambled sentences is still not known. However, studies have suggested that scrambled sentences are associated with certain functions, instead of being a free variation of their canonical counterparts.

Yamashita (2002) examined the functions of scrambled sentences by analyzing a variety of texts with varying genre and formality, and found two properties that the scrambled constituents had. The first one was “heaviness” of the constituent. In the 14 (out of 19) scrambled sentences in the data, the scrambled constituents were “heavy” in that they contained a subordinate clause and consisted of several content words. This “long before short” tendency was also observed experimentally in Yamashita and Chang (2001). The second property of the scrambled constituents was “referentiality”. In the five scrambled sentences in the data, the scrambled constituents had a determiner or an anaphor that referred to some elements in the preceding context. “Heaviness” and “referentiality” accounted for 95% of the scrambled sentences in the data. Kondo and Yamashita (2011) also confirmed the effect of “heaviness” and “referentiality” in their analysis of spoken corpus data.

Yamashita (2002) further investigated whether the scrambling structure might be associated with two discourse principles, “given before new” (e.g., Halliday, 1967) and “change of topic” (e.g., Fletcher, 1985), which are considered to facilitate comprehension. However, Yamashita found no reliable connection between such discourse principles and the use of scrambling, and thus proposed that the motivation for scrambling might be production-based rather than comprehension-based. According to Yamashita, one possibility is that scrambling contributes to the efficiency of the speaker’s sentence production. Scrambling of a heavy constituent (e.g., subordinate clause) allows the speakers to process one clause at a time, resulting in a reduced psychological cost. The other possible production-based motivation of scrambling is that, due to the prominence or the accessibility of a particular constituent, the speaker might want to say the constituent first.[3]

[3] While Yamashita (2002) maintains that the use of scrambling is motivated by production rather than comprehension, scrambling motivated by “heaviness” and “referentiality” may also contribute to ease of comprehension (T. G. Bever, personal communication; Mazuka, Itoh, & Kondo, 2002).

An experimental study by V. Ferreira and Yoshita (2003) supports the accessibility account of scrambling. Below are examples of their material.

(1.8) *Questions:*

a. Okusan-ga otetsudaisan-ni kanshashiteita. Sorekara doushita?
housewife-Nom housekeeper-Dat was grateful then what
'The housewife was grateful to the housekeeper. What happened next?'

b. Okusan-ga meidosan-ni kanshashiteita. Sorekara doushita?
housewife-Nom housemaid-Dat was grateful then what
'The housewife was grateful to the housemaid. What happened next?'

(1.9) *Targets:*

a. Okusan-ga otetsudaisan-ni purezento-o okutta.
housewife-Nom housekeeper-Dat present-Acc gave
'The housewife gave the housekeeper a present.'

b. Okusan-ga purezento-o otetsudaisan-ni okutta.
housewife-Nom present-Acc housekeeper-Dat gave
'The housewife gave a present to the housekeeper.'

The participants were instructed to match the questions and the targets first and were also asked to memorize the targets. Then, they listened to the questions and recalled the matching targets. The results indicated that the participants tended to place the word that appeared in the question (the given word) before the word that did not (the new word) when they responded to the question. It was also found that the effect was more robust when the question and the target shared the same word ('housekeeper' vs. 'housemaid'). Thus the results suggest that the choice of word order is sensitive to the accessibility of the arguments.

1.2.3 Frequency of scrambling

While scrambling of NPs is quite common in Japanese, studies indicate that the use of the OSV word order in particular is far less frequent than its canonical counterpart. For instance, a survey by Kokuritsu Kokugo Kenkyujo (National Institute of Japanese Language and Linguistics, 1964) found that, in the magazine articles analyzed, 467 sentences were in canonical word order (SOV) while only 27 sentences had a scrambled word order. In an analysis of newspaper articles, Kuno (1973a) found that canonically ordered sentences (SOV) were 17 times more likely to occur than OSV sentences. Yamashita (2002) analyzed four types of magazine articles with varying degrees of formality. It was found that, of the 2,635 sentences analyzed, there were only 19 scrambled sentences (less than one percent). Infrequency of OSV sentences in those analyses, however, may be attributed to the fact that the sources were all in heavily edited contexts, where OSV sentences might have been intentionally suppressed.

Scrambled sentences may be used more frequently in spoken contexts than in written contexts. Kondo and Yamashita (2011) analyzed the spoken data in the Corpus of Spontaneous Japanese (National Institute of Japanese Language & National Institute of Information and Communications Technology, 2004). Three types of sentences were analyzed: transitive sentences, ditransitive sentences, and transitive sentences with a locative phrase.[4] Their frequency analyses indicated that the ratios of the scrambled word order were 6.5%, 31.7%, and 49.1% for the transitives, ditransitives, and transitives with a locative phrase, respectively. Kondo and Yamashita attribute the lower frequency of scrambled transitive sentences (OSV) to the prominence of the functional role that the subject plays in Japanese transitive sentences. The subject in a transitive sentence is often omitted in Japanese, and the overt realization of the subject tends to be specifically motivated, such as exhaustive (i.e., “X and only X”)

[4] Note that the scrambling of the transitive sentences with a locative phrase in Kondo and Yamashita (2011) was the fronting of the locative NP before the accusative NP (i.e., S O Locative V → S Locative O V). It is different from the scrambling of the locative sentences examined in Task 1 of the present study, which was the fronting of the locative NP before the nominative NP (i.e., S Locative V → Locative S V).

interpretation of the subject (Kuno, 1973b).[5] Such prominence of the subject might motivate speakers to utter the subject first, resulting in a smaller ratio of OSV sentences.

It is probably worth noting that scrambling might be seldom taught or practiced in Japanese language instruction. One Japanese textbook which is used widely at colleges in the United States (*Genki I & II;* Banno, Ohno, Sakane, Shinagawa, & Tokashiki, 1999) spares only about a third of a page to illustrate the freer word-order phenomenon. In another commonly used textbook (*Nakama 1 & 2 2nd Edition;* Hatasa, Hatasa, & Makino, 2009), the author was unable to identify any mention of the flexible word order of NPs. The possible lack of overt instruction, coupled with the relative infrequency of scrambled sentences, may leave many L2 learners unaware that scrambling is permissible in Japanese.

1.3 The present study

One objective of the present study was to investigate how Japanese sentences in the OSV order are produced and comprehended by L2 learners. As will be reviewed in the subsequent chapters, studies with Japanese-speaking children have found that they go through a developmental stage during which they consistently misunderstand scrambled sentences, taking the first NP in OSV sentences to be the subject. Moreover, it has been found that NS adults experience slowdowns in reading and comprehending scrambled sentences. On the other hand, investigations into the processing of Japanese scrambled sentences by L2 learners have been scarce, and it is not entirely clear how sentences with alternative word orders are processed by L2 learners of Japanese. The present study attempted to investigate the source of difficulty

[5] The preference of omitting the subject among NSs of Japanese has been observed experimentally. Miyamoto and Nakamura (2005) reports that, when asked to complete a sentence starting with an object (NP-*o*), the NSs omitted a subject in 210 instances while they included a subject in 39 instances.

during L2 processing of scrambled sentences as well as the difference between L1 and L2 processing of such sentences.

Another objective of the present study was more pedagogical in nature. There has been a growing body of research that has attempted to apply psycholinguistic techniques to language learning, and the results have been generally positive. In this study, methods involving *syntactic persistence* (a tendency for language users to reuse a previously produced or comprehended sentence structure for the processing of subsequent sentences) were employed to examine whether such methods might be effective in facilitating the production of OSV sentences among L2 learners.

1.3.1 Research questions

The present study attempted to examine the following three main questions.

Question 1: Do L2 learners experience difficulty in producing OSV sentences as compared to SOV sentences? If so, is the source of difficulty the lack of grammatical knowledge or a processing problem? There have been several studies that investigated the comprehension of scrambled sentences by L2 learners. On the other hand, studies on production are scarce, and it is not known whether L2 learners experience difficulty when prompted to produce scrambled sentences. Moreover, while previous studies have found lower comprehension accuracy associated with scrambled sentences, it is not clear whether such lower accuracy is due to the lack of knowledge about scrambling (the knowledge that scrambling is permissible in Japanese and/or the knowledge that word order is not a reliable cue in thematic role assignment in Japanese) or due to a processing problem. The present study attempted to assess L2 learners' grammatical knowledge about case markers and their production performance of OSV sentences through a fill-in-the-blank task (Task 1) and a picture description task (Task 2). The data from the two tasks were further compared to determine whether L2 learners' difficulty in producing OSV sentences might be derived from their lack of grammatical knowledge or processing strategies.

Question 2: Is the comprehension process of OSV sentences different between L2 learners and NSs? Studies that examined the comprehension of scrambled sentences by NSs have provided evidence that there is a psychological cost in processing scrambled sentences. That is, it takes NSs longer to read and comprehend scrambled sentences than canonical ones. The present study attempted to examine, through a sentence correctness decision task (the pilot study) and a self-paced reading task (Task 3), whether L2 learners would also experience such slowdowns, and if so, whether the pattern of slowdown would be similar or different between L2 learners and NSs. Through the results, this study also aimed at examining the claim that L2 sentence processing might be qualitatively different from its L1 counterpart (Shallow Structure Hypothesis; Clahsen & Felser, 2006).

Question 3: Is syntactic persistence effective in facilitating the production of OSV sentences? Does it improve the use of postpositional particles among L2 learners? As discussed briefly above, studies that investigated the L2 sentence production using syntactic persistence have yielded promising results, showing a positive contribution of syntactic persistence in the acquisition of L2 grammatical structures. The present study examined whether syntactic persistence tasks might increase the production of OSV sentences to explore the possibility of using such a method in scaffolding the learning of scrambling (Tasks 4 & 6). Additionally, the results of a fill-in-the-blank task (Task 5) and a syntactic persistence task (Task 6) were compared to examine whether the syntactic persistence task would be effective in improving the correct use of postpositional particles.

In the examination of the main questions above, this study also attempted to investigate the following issues.

Question 4: What is the relationship between the grammatical knowledge/processing of the scrambled sentences and L2 learners' general proficiency in Japanese? Are more proficient learners better at processing OSV sentences? Some studies have suggested that L2 learners' processing strategies in comprehending Japanese sentences become more similar to native speakers' (prioritizing case marking information over word order for thematic role assignment) as they acquire higher general proficiency in the target language. Other studies, however, have found that comprehension and production of scrambled sentences did not necessarily improve with the learners' proficiency in Japanese. Thus, it would be fruitful to examine whether proficiency might correlate with grammatical knowledge and processing strategies for scrambled sentences. The present study attempted to examine the issue by observing the comprehension and production of OSV sentences by different proficiency groups.

Question 5: What are the effects of "reversibility" on L2 learners' processing of OSV sentences? Do L2 learners integrate the animacy of NPs during their processing of scrambled sentences? Although L2 studies on the processing of scrambling have observed difficulty associated with the scrambling structure, whether animacy of the NPs might interact with such difficulty has not been studied. The present study included two sentence types to examine the point: "reversible" sentences (those with animate subject and animate object) and "non-reversible" sentences (those with animate subject and inanimate object). In reversible sentences, because both nominative and accusative NPs are animate, the only information signaling that the sentence is scrambled are the case markers. In non-reversible sentences, on the other hand, the "inanimate-animate" NP order would also signal that the sentence is scrambled because it is less likely, from our general knowledge, that an inanimate subject would do something to an animate object. Thus, scrambling is possibly more salient for non-reversible sentences. This study observed whether the two sentence types might interact with L2 learner's processing of OSV sentences.

1.3.2 Overview of the methods

1.3.2.1 Participants

Unless otherwise noted, the participants in the tasks reported in the present study were 54 non-native speakers (NNS) and 20 native speakers (NS) of Japanese.

Non-native speaker participants

Fifty-four non-native speakers of Japanese (18 females and 36 males) participated in the study.[6] All of the NNS participants except one (NNS54) were recruited from within the University of Arizona community. Due to the concern that participants' L1 might influence their

[6] One male NNS participant (NNS20) was unable to participate in the second session of the study, and therefore, the data from this participant were not included in the analyses.

processing of Japanese sentences, only those learners of Japanese whose L1 is English were recruited in the present study.[7]

All of the NNS participants but one (NNS22) had completed at least four semesters of formal classroom instruction in Japanese. The mean and median of the length of their Japanese studies were 42.8 months and 25 months, respectively. Their ages ranged from 19 to 64 years old, but the majority of them were in their 20s (mean: 24.1, median: 22). Twenty-five of them had stayed in Japan for more than one month. Their self-rating of overall proficiency in Japanese on a 7-point Likert scale ("1" being "beginner" and "7" being "native-like") ranged from 1 to 6 (mean: 4.04, median: 4).

In order to assess the participants' proficiency in Japanese, Versions B and A of the Simple Performance-Oriented Test (SPOT) (Ford-Niwa & Kobayashi, 1999; Kobayashi 1997, 1998) developed by the International Student Center at the University of Tsukuba, were administered individually at the beginning of the first study session.[8] The SPOT test consists of isolated sentences printed on answer sheets and an audio recording. The sentences on the answer sheets are written in Japanese orthography, but each sentence is missing a grammatical element or a part of grammatical element whose location is specified with parentheses. The test-taker's task is to fill in the parentheses with one *hiragana* while they listen to the audio recording of the sentences. For the purpose of assessing the test-taker's online and automatic processing of the language, the sentences are recorded rather fast and there are only two-second pauses between the sentences. There are 60 items in Version B of SPOT and 65 items in Version A. Because Version B is the easier of the two, the participants in the present study took Version B first

[7] Five of the NNS participants are bilingual speakers of English and Spanish.
[8] As of August 2014, an Internet-based version of SPOT, which consists of 90 test items, is available on the website of the International Student Center at the University of Tsukuba (http://www.ttbj.jp/).

followed by Version A, as recommended by the test developers. In the present study, the participants listened to the recording from the speaker of a laptop computer.

Based on the sum of the scores from the two versions of the test, the participants were placed into three groups: Low-Intermediate (LI), High-Intermediate (HI), and Advanced (ADV). Twenty, fourteen, and twenty participants were placed into the LI, HI, and ADV groups, respectively. Out of the 125 possible points from the two versions of SPOT, the scores of the LI group participants ranged from 22 to 72, the HI group from 80 to 96, and the ADV group from 101 to 121.

While the validity and reliability of SPOT has been extensively studied and established (e.g., Kobayashi, 1997, 1998), the point-biserial correlations of each test item were calculated for the purpose of examining the validity of the grouping of the participants based on their SPOT scores. Point-biserial correlation measures the correlation between the scores that individuals received on a given test item and the total score that the same individuals received on the test. A low point-biserial correlation coefficient indicates that the item has low discriminating power to separate high-scoring individuals and low-scoring individuals. According to Fulcher and Davidson (2007), items with point-biserial correlation of 0.25 are generally considered acceptable, and therefore, this value was chosen as the cutoff in this item analysis. There were 12 items whose point-biserial correlation coefficients were below 0.25 (item numbers 1, 7, 31, 54, 57 of SPOT Version B, and item numbers 1, 9, 19, 51, 54, 55, 65 of SPOT Version A). The adjusted scores were calculated for each participant after eliminating the 12 items. However, there were no changes in the proficiency groupings (that is, there was no crossover of scores between the highest-scoring participant in a lower proficiency group and the lowest-scoring participant in a higher proficiency group either before or after the application of the adjusted scores). The adjusted scores of the LI group participants ranged from 15 to 63, the HI group from 73 to 87, and the ADV group from 91 to 111. Table 1.1 summarizes the background of the participants in each proficiency group, and Table 1.2 provides a summary of the proficiency scores of the participants in each group. See Appendices 1A and 1B for more detailed information on the individual NNS participants' backgrounds and their proficiency scores.

Table 1.1: Summary of NNS participants' background

	LI		HI		ADV	
Number of participants (female, male)	6	14	5	9	7	13
Age (mean, median)	22.6	21.5	22.9	22.5	26.6	23
Time studied Japanese in month (mean, median)	25.6	22	38.2	30.5	63.1	45
Time spent in Japan in month (mean, median)	0.6	0	6	0	16.3	4.8
Highest J. class completed in semester (range)	4th-6th		2nd-8th		4th-8th	
Frequency of Japanese use (mean, median)	4.5	5	4.4	4.5	5.1	5

Note: The participants were asked to indicate how often they used Japanese in the past 6 months using a 7-point Likert scale ("1" being "not at all" and "7" being "very frequently") for "Frequency of Japanese use".

Table 1.2: Ranges and means of the SPOT scores in percent by proficiency group

	LI		HI		ADV	
SPOT B	21.7-71.7	(52.4)	70.0-90.0	(79.3)	85.0-100.0	(93.1)
SPOT A	13.8-50.8	(37.3)	53.8-67.7	(62.5)	72.3-93.8	(83.3)
SPOT B+A	17.6-57.6	(44.6)	64.0-76.8	(70.6)	80.8-96.8	(88.0)
Adjusted	13.3-55.8	(42.6)	64.6-77.0	(70.7)	80.5-98.2	(89.2)

Note: The numbers in parentheses are means. "Adjusted" stands for the scores after the 12 items below the point-biserial correlation of 0.25 were eliminated.

Native speaker participants

Twenty native speakers of Japanese (9 females and 11males) also participated in the study. As with the NNS participants, they were recruited from within the University of Arizona community. They all completed their high school education in Japan. The NS participants' ages ranged from 19 to 36 years old (mean: 26.7, median: 26.5), and the length of their overseas experiences (out of Japan) ranged from 2 months to 10 years (mean: 46.9 months, median: 48.5 months).

1.3.2.2 Tasks

The results of seven tasks are reported in the present study. Table 1.3 below presents the type, main purpose, and participants of each task.[9]

Table 1.3: Type, purpose, and participants of the seven tasks in the present study

	Type	Main purpose	Participants
Task 1	Fill-in-the-blank	Assessing grammatical knowledge about particles	NNS only
Task 2	Picture description	Assessing production performance of OSV sentences	NNS & NS
Task 3	Self-paced reading	Observing pattern of slowdown in reading OSV sentences	NNS & NS
Task 4	Syntactic persistence (transitive sentences as primes)	Observing the syntactic persistence effect for OSV sentences	NNS & NS
Pilot Study	Sentence correctness decision	Observing whether slowdowns take place in comprehending OSV sentences	24 NNSs, NS
Task 5	Fill-in-the-blank	Assessing the grammatical knowledge about particles	A subgroup of NNS (30)
Task 6	Syntactic persistence (questions as primes)	Observing the syntactic persistence effect for OSV sentences	A subgroup of NNS (30) & NS (16)

1.3.2.3 Order of participation

Participation for Tasks 1 through 4 (and for the pilot study for the NS participants) were divided into two study sessions. The study sessions were scheduled individually for each participant, and there was typically one week between the two sessions. For the first study session, the NNS participants completed the consent process followed by a language background

[9] None of the 24 NNS participants in the pilot study were among those who took part in the other tasks in the present study (Tasks 1-6), while the 20 NS participants who took part in the other tasks in the present study also participated in the pilot study.

questionnaire, which asked their experiences with second languages. The information obtained through the questionnaire is presented in Appendices 1A and 1B. The NNS participants then took the SPOT test, and completed Tasks 1 and 2 during the first session. The NS participants also completed the consent process and a language background questionnaire first. Then they worked on Task 2 and the pilot study.

For the second study session, both NNS and NS participants completed Tasks 3 and 4. At the end of the second session, more than half of the participants also went through a debriefing interview during which the purposes of the present study were explained and the participants' impressions and experiences with scrambled sentences were asked.

A subgroup of the individuals who took part in Tasks 1 through 4 (30 NNSs and 16 NSs) participated in Tasks 5 and 6. About half of the participants in these two tasks had gone through the debriefing interview for the previous tasks, and the other half worked on Tasks 5 and 6 as the third study session before the debriefing interview. In order to minimize the effect of the debriefing interview for those who had already gone through the procedure, the participants were invited to Tasks 5 and 6 only if there had been more than two months since they took the debriefing interview.

1.4 Organization of chapters

The organization of the present study is as follows. Chapter 2 reviews previous studies on L1 and L2 acquisition of the scrambling structure, and reports the results of Tasks 1 and 2, which were aimed at investigating L2 production of OSV sentences. Chapter 3 investigates comprehension processes of OSV sentences by L2 learners through Task 3 as well as the pilot study. Chapter 4 examines the effectiveness of syntactic persistence in facilitating the production of OSV sentences as well as improving the use of postpositional particles, through Task 4 and the follow-up tasks (Tasks 5 & 6). Chapter 5 summarizes the findings of this study, and provides provisional answers to the research questions based on the findings. It also discusses pedagogical implications of this study, and suggests some directions for future research.

CHAPTER 2: THE RELATION OF GRAMMATICAL KNOWLEDGE AND PERFORMANCE IN L2 LEARNERS' PRODUCTION OF SCRAMBLED SENTENCES

2.1 Introduction

The chapter will report the results of Tasks 1 and 2. Task 1 (fill-in-the-blank) attempted to assess L2 Japanese learners' grammatical knowledge as the basis for processing scrambled sentences. Task 2 (picture description) attempted to assess their production of scrambled sentences. English does not allow scrambling of the accusative NP before the nominative NP within the same clause except for the rather infrequent instances of topicalization, whereas Japanese allows this grammatically. When L1 English learners of Japanese are prompted to produce Japanese sentences in a scrambled word order, will there be more errors as compared to their production of sentences in a canonical order? If so, are the errors due to a lack of the grammatical knowledge or to processing difficulty during production? Does their grammatical knowledge ensure the correct production of scrambled sentences?

The chapter will first briefly review two concepts (canonical template and reversibility) that are most relevant to the present study. It will review studies that have examined the acquisition of scrambling structure by L1 Japanese children, as these L1 studies provide important implications for the acquisition of scrambling structure by L2 learners. A review of previous L2 acquisition studies and a discussion of the present study will follow.

2.2 Two concepts that are most relevant to the present study

2.2.1 Canonical template

The first concept most relevant to the present study is the canonical template, which refers to surface forms in a given language that can be used for heuristic sentence processing. Extensive research on sentence comprehension by L1 English children has found that they go through a developmental stage during which their comprehension accuracy of passive and object-cleft sentences systematically declines. Bever (1970) suggested that this comprehension decline is due to the development of a perceptual strategy now commonly known as the NVN strategy. The N(oun)V(erb)N(oun) sequence, which corresponds with the agent, action, and patient of a sentence, is quantitatively most frequent in English. L1 English children around age four appear to overuse this NVN template in their sentence comprehension, interpreting the first NP of a sentence as the agent regardless of the actual grammatical structure of the sentence. This overgeneralization of the NVN template seems to cause the decline in comprehension accuracy of passive and object cleft sentences because the first NP in such sentences is the patient rather than the agent.

The development of such templates for sentence processing appears to be observed cross-linguistically. Slobin and Bever (1982) studied the comprehension of simple transitive sentences by children acquiring different L1s (English, Italian, Serbo-Croatian, and Turkish), and their data provided evidence that children, regardless of the language they are acquiring, develop and use a canonical sentence template that conforms to the typical features (e.g., word order, inflection, and prosody) of the input language. In English and Italian, the grammatical relations of the NPs in a sentence are determined mostly by word order, but Italian makes more frequent use of stress contrasts to determine the grammatical relations. It was found that the Italian children made greater use of the information provided by stress than the English children while both groups of children were more consistent in their use of word order information. Turkish, on the other hand, is an inflectional language with more flexibility in word order, and therefore, the grammatical relations of the NPs in a Turkish sentence is determined by inflection rather than word order. Consequently, the Turkish children exhibited a more consistent use of inflection information than word order. As for the use of word order strategy, the researchers observed that it was most prominent around 3;6 (year ; month) across all L1 groups, which corresponds with Bever's (1970) observation.

While adults generally do not blindly apply the canonical template to sentence processing, taking advantage of grammatical knowledge to determine grammatical relations among NPs, the schema that has been developed based on the typical features of any given language seems to play an important role in adult sentence processing as well. Bates, McNew, MacWhinney, Devescovi, and Smith (1982) examined the interpretation of simple sentences in varying word order by adult speakers of English and Italian. Italian allows more flexibility in word order than English, especially in informal discourse, and therefore, the researchers suspected that word order might play a less important role in the determination of grammatical relations in Italian. The within-subject variables in addition to word order were stress, animacy, and topicalization, and the participants' task was to identify which of the two nouns in a sentence they had just heard was the agent.[10] The results indicated that the L1 Italian adults relied more on semantic information (i.e., animacy) in the interpretation of the sentences while L1 English adults relied on word order. It was also found that the Italian adults made greater use of the stress and topicalization information. Bates et al. (1982) thus demonstrated that a canonical sentence

---

[10] The topic manipulation included a prompt such as the following before the presentation of the sentence: "This is a (cow); now we're going to talk about this (cow)" (Bates et al, 1982: 264).

schema is developed based on the features of the input language and that adults apply such a schema for their sentence processing. More recently, Townsend and Bever (2001) proposed a hybrid model of sentence comprehension (the analysis by synthesis model) which integrated a probabilistic analysis based on the canonical schema with full syntactic parsing.

2.2.2 Reversibility

Reversibility refers to the interchangeability of the subject and the object NPs in a sentence, and it has been frequently included as a variable in sentence processing studies for the purpose of investigating the effects of semantic strategy. Reversible sentences refer to those in which interchanging the two NPs is possible without degrading the acceptability of the sentence, e.g., *The boy kicked the girl → The girl kicked the boy*. On the other hand, non-reversible sentences refer to those in which switching of the subject and object NPs result in significant degradation. According to Forster and Olbrei (1972), there are two types of non-reversible sentences. The first type is logically non-reversible sentences in which switching the two NPs results in an anomalous sentence, e.g., *The man ate the pizza → The pizza ate the man*. The other type is pragmatically non-reversible sentences in which the reversing of the NPs makes the sentence implausible, e.g., *The mother fed the baby → The baby fed the mother*.

It is generally considered that non-reversible sentences are easier to process than reversible ones, especially in the passive form (e.g., Baldie, 1976; Harris, 1976; Hayhurst, 1967; Slobin, 1966; Turner & Rommetveit, 1967). This is so because even when the positions of the subject and object NPs are switched in non-reversible sentences, as in passive constructions, the semantics of the NPs inform which NP is the agent of the action. Slobin (1966) examined the processing of kernel, negative, passive, and passive negative sentences with reversible and non-reversible contrasts within each sentence type. The non-reversible sentences used in the study were of the "logical" type. The participants were four age groups of children (6, 8, 10, and 12 years old) and adults, and their task was to judge whether the sentence they heard and the picture they saw matched. The results indicated that the non-reversible sentences were responded to faster than reversible ones across all age groups. It was also found that the non-reversible sentences were especially facilitative with respect to passive and passive negative sentences. That is, while the differences in reaction times between the kernel and passive sentences as well as between the negative and negative passive sentences were significant for the reversible sentences, such differences were not observed for the non-reversible sentences. Turner and Rommetveit (1967) tested the imitation, comprehension, and production of active/passive and reversible/non-reversible sentences by children of five age groups. Similarly to Slobin (1966), their results indicated that non-reversible sentences were constantly easier than reversible counterparts for children to imitate, comprehend, and produce.[11]

On the other hand, Forster and Olbrei (1972) provided evidence that reversible sentences are not necessarily more difficult to process than non-reversible sentences, at least for adults who have developed full-fledged syntactic processing. Using carefully controlled items in terms of plausibility, they tested the processing of active/passive and reversible/non-reversible sentences. They used sentence correctness decision tasks, in which the participant is presented with a sentence and makes a decision as to whether the sentence is correct as quickly as possible, as well as an RSVP (rapid serial visual presentation) task in which each word of a sentence is

[11] Bever (1970) presents evidence that implausible sentences such as "The dog pats the mother." (p. 306) are comprehended better by two-year-old children better than three-year-old children, suggesting that the use of plausibility as a perceptual strategy depends on experience. As observed, Bever suggests that L1 English children go through a developmental stage in which they heavily depend on the NVN strategy (around the age 4) shortly after they utilize plausibility for sentence comprehension.

presented very rapidly in a serial manner and the participant names as many words as possible from the presentation. The results of the sentence correctness decision tasks indicated that the difference in reaction latencies between the reversible and non-reversible sentences was minimal. More importantly, there was no significant interaction of reversibility and voice, which indicates that non-reversibility did not facilitate the processing of the passives. The result of the RSVP task did not observe any difference between the reversible and non-reversible sentences. Based on the results, Forster and Olbrei suggested that the reversibility effect observed in earlier studies might be attributable to the items and the procedures employed, and that the syntactic analysis itself might be independent of the meaning of the sentence.

F. Ferreira (2003) also investigated the influence of reversibility on the comprehension of active and passive sentences (Experiment 1). There were three types of stimuli in terms of reversibility: *biased reversible* (e.g., the dog bit the man/the man bit the dog), *non-reversible* (e.g., the mouse ate the cheese/the cheese ate the mouse), and *symmetrical reversible* (e.g., the woman visited the man/the man visited the woman). The participants (college-age native speakers of English) listened to active and passive versions of those sentences and identified the agent or patient/theme. The results indicated that passives were more difficult to understand than actives regardless of the reversibility sentence types. It was also found that plausible sentences (e.g., the dog bit the man) were comprehended more accurately than implausible ones (e.g., the man bit the dog). Because another experiment (Experiment 3) also observed that object-clefts (e.g., it was the dog the man bit), whose interpretation do not conform to the NVN template just as passives do not, were more difficult to comprehend than subject-clefts (e.g. it was the man who bit the dog), Ferreira suggested that sentence processing involves heuristics such as the

NVN template and semantic association ("good enough" processing) in addition to more elaborate syntactic algorithms.

For the processing of Japanese scrambled sentences, the word order template predicts that scrambled sentences are more difficult or psychologically more costly to process because the canonical Japanese word order is SOV. This prediction has been confirmed through quite a few experimental studies, some of which will be reviewed below and in the next chapter. If we assume that the semantics of nouns interact with sentence production, the canonical template might predict that reversible OSV sentences would cause more errors than non-reversible sentences (the logical type, at least) in production. Because Japanese sentences generally prefer animate subjects to inanimate ones, and the first NP in reversible scrambled sentences is animate, users of Japanese may be more prone to errors based on the SOV template when they are instructed to produce reversible OSV sentences, at least in an experimental setting. This point will be examined in Task 2 of the present study.

2.3 Literature review

2.3.1 Processing of OSV sentences by L1 Japanese children

The first inquiries into the processing of Japanese scrambling were conducted with NS children in the context of first language acquisition. Such studies were inspired by Bever's (1970) theory of perceptual strategy in sentence comprehension, and were conducted to verify the word order strategy empirically among NS Japanese children.

In a pioneering study, Hayashibe (1975) used two act-out procedures to examine the comprehension of active SOV and OSV sentences by children between 3;0 and 5;11. In the first experiment, the experimenter read aloud three words (two nouns without case markers and a verb), and children were instructed to act out the meaning of the sentences using toys. The result revealed that four and five-year-old children showed a strong tendency to interpret the first noun as the agent, while three-year-old children showed a more "egocentric" interpretation (i.e., takes himself as the agent of the sentence). In the second experiment in which children were instructed to act out based on case-marked SOV and OSV sentences, it was found that there was a period during which children relied heavily on word order before case-marking particles are used reliably as comprehension cues. Based on the findings, Hayashibe concluded that SOV word order is dominant to OSV, and that the word-order cue is acquired before case-marking particles.

Sano (1977) examined the comprehension of active, passive, and cleft sentences by children between 3;3 and 6;8 using act-out and imitation tasks. The materials included both canonical and scrambled sentence types. The results of the comprehension task suggested that children's ability to comprehend canonical sentences developed earlier than their ability to comprehend scrambled sentences for both active and passive sentences. For cleft sentences, the children showed a strong tendency to interpret the clefted NP as the agent.

In the imitation task, the children tended to interchange the particles in imitating OSV sentences, so that the resultant sentences would have the canonical SOV word order. The results of the imitation task also showed that, when simple active sentences with omitted particles were presented, the children displayed a strong tendency to supply the –*o* particle after the second NP, the position right before the verb. On the other hand, when simple and cleft passive sentences were presented, the children often supplied the –*ni* particle at the position immediately preceding the verb. Based on the observation, Sano suggested that a particle in the position immediately preceding the verb may play a significant role in children's sentence comprehension.

Hakuta (1982) examined the interaction of case markers and word order in comprehension and production by Japanese children. The following four types of simple active and passive sentences were tested:

(2.1)	Active/SOV:	Agent-ga Patient-o active-V.
	Active/OSV:	Patient-o Agent-ga active-V.
	Passive/SOV:	Patient-ga Agent-ni passive-V.
	Passive/OSV:	Agent-ni Patient-ga passive-V.

If Japanese children rely on the NNV (agent, patient, action) word-order strategy, analogous to what L1 English children do, their performance on Active/SOV and Passive/OSV sentences should be better than the performance on Active/OSV and Passive/SOV sentences. If, on the other hand, children interpret case markers properly, their performance on canonical and scrambled sentences should not differ.

In the sentence comprehension experiment (using the act-out task), Hakuta's participants (children between 2;3 and 6;2) performed on the Active/SOV sentences significantly better than on other sentence types. For sequences without case-markings, the children tended to choose the first noun as the agent of the action, but such a tendency was strongest for the NVN sequence and weakest for NNV sequence, which corresponds to the canonical Japanese form.

Hakuta further tested sentence production by children using a picture description task. The results showed that the sentences produced were predominantly the Active/SOV type. Among the 576 opportunities, there were only 13 instances of OSV active, 14 instances of SOV passive, and no instances of OSV passive sentences. The particle *–ga* was supplied far more frequently than *–o*, suggesting that children acquire control over *–ga* before *–o*. The results of the imitation task also showed a preference for SOV sentences and a strong bias for matching the first NP and the *–ga* marker. Children encountered difficulty imitating the OSV sentences correctly; the errors mostly involved particles, where the order of the nouns was retained but the first noun was marked by *–ga*. Overall, therefore, Hakuta's study showed that Japanese children take advantage of canonical form as do English-speaking children, but "Japanese children's requirement for overgeneralization of the canonical form is more specific" (p. 73). That is, while English-speaking children may rely solely on word order, Japanese children may employ the presence of the *–ga* marker on the first NP as a cue in addition to the canonical NNV word order. Goto (1989), using picture description and verification tasks, also observed that Japanese children integrated the word-order and case-marker cues for the processing of simple transitive sentences, and that the case-marker cue becomes prioritized over the word-order cue after the age of five.

While experimental studies have generally observed that the acquisition of the scrambling construction by Japanese children is rather late (i.e., up to five years old in Hayashibe, 1975), studies based on naturally occurring data indicate that scrambled sentences are produced earlier (although "production" cannot necessarily be equated with "acquisition"). In a longitudinal study of a girl from birth to three years old, Miyahara (1974) observed that the use of the topic marker *–wa* and the subject marker *–ga* started to appear around 1;8 and the object marker *–o* slightly later around 1;11. Miyahara noted that increased flexibility of word order was observed only after the *–o* particle was acquired. Iwatate (1981) observed 16 occurrences of SOV sentences and 10 OSV sentences by a boy during the 2;2-2;8 period in naturally occurring data recorded over 51 hours. In the same study, Iwatate observed 79 SOV sentences and 10 OSV sentences in 42 hours of audio-recorded spontaneous utterances of five children between 2;5 and 3;9.

Based on Iwatate (1981), who observed production of scrambled sentences by two-year-old children, Otsu (1994b) suspected that the grammatical knowledge about scrambling should develop earlier (by the age of three) than the previous experimental studies had suggested (e.g., Hayashibe, 1975), and that the results of the experimental studies might not have reflected the children's grammatical knowledge. In other words, the children in the experimental studies might not have been able to make use of their grammatical knowledge and relied on the NNV strategy instead because of the way the scrambled sentences were presented to them. While OSV sentences were presented in isolation in the previous studies, Otsu used materials such as the following:

(2.2) Kooen-ni ahirusan-ga imashita.
park-in duck-Nom was
'There was a duck in the park.'

Sono ahirusan-o kamesan-ga oshimashita.
the duck-Acc turtle-Nom pushed
'A turtle pushed the duck.'

The first sentence provided context to the second sentence, making the scrambling of the second sentence more natural. Otsu presented such sentences to three and four-year-old children and compared their comprehension of scrambled sentences with that of children in the control group who received scrambled sentences without contextual cues. The results showed that, while the control group showed errors in comprehending scrambled sentences as did the children in previous studies, the children in the experimental group comprehended scrambled sentences as accurately as canonical sentences.

Otsu (1994a) also demonstrated that three and four-year-old children were able to produce scrambled sentences in an experimental setting. The task employed in this study was picture description. After making sure that the participants were able to correctly name the objects in the picture, the experimenter asked the following question:

(2.3) Kono e- nituite ohanasi-sitekureru? Mazu, X-de hazimetene?
this picture about talk for me first with begin
'Can you tell me about this picture? First, can you begin with X?'

The result showed that, when prompted to start with the object, three out of 10 three-year-olds and eight out of 10 four-year-olds used the OSV structure correctly at least once, and none of the participants produced an ungrammatical sentence.

Murasugi (2000) and Murasugi and Kawamura (2004) also provide evidence that scrambling is acquired earlier than generally assumed. They tested the comprehension of canonical, scrambled, and passive sentences among children between ages two and six (two and four in Murasugi, 2000) using an act-out task. The result showed that the participants, even the two and three-year-olds, responded to scrambled sentences as accurately as canonical sentences. However, the response accuracy of passives was generally lower than that of canonical and scrambled sentences.

Neither Murasugi (2000) nor Murasugi and Kawamura (2004) included discourse context such as that used in Otsu (1994b), yet the participants in the studies did not seem to have difficulty responding to scrambled sentences unlike previous experimental studies. The researchers attributed the difference to the inclusion of the passive sentences in their materials. They speculated that the inclusion of passives forced the children to pay more attention to the case markers and the theta-roles that they assign, which in turn resulted in a higher comprehension accuracy of the scrambled sentences.[12]

[12] The precise reasons for the high comprehension accuracy with scrambled sentences in these studies are unknown, as previous studies such as Sano (1977) and Hakuta (1982) included sentence types other than active canonical/scrambled yet the participants performed poorly with scrambled sentences. It might be the case that it was easier for the participants in Murasugi's study to develop strategies to complete the task, as there were only three sentence types in the material. There were 13 sentence types (Stimulus I) and 18 sentence types (Stimulus II) in Sano (1977), and there were four in Hakuta (1982).

Murasugi and Kawamura (2004) further tested L1 Japanese children's knowledge of the reconstruction property of scrambling, using materials such as the following:

(2.4) a. Ahiru-ga usi-o [zibun-no niwa-de] oikaketa.
duck-Nom cow-Acc self-Gen garden-at chased
'The duck chased the cow at the garden of himself.'

b. Usi-o$_i$ [zibun-no niwa-de]$_j$ ahiru-ga t_j t_i oikaketa.
cow-Acc self-Gen garden-at duck-Nom chased
'The cow, at the garden of himself, the duck chased.'

As discussed in Saito (1985) and others, scrambling in Japanese exhibits the reconstruction property. That is, the scrambled element may be interpreted at the position where it originated. The anaphor *zibun* requires a c-commanding antecedent for interpretation. In (2.4a), the subject *Ahiru-ga* properly c-commands *zibun,* and therefore, *zibun-no niwa* is interpreted as 'the duck's garden'. In (2.4b), on the other hand, this c-command relationship is destroyed, yet the interpretation of (2.4b) is essentially the same as that of (2.4a). This suggests that the locative PP needs to be moved back to its underlying position before scrambling for correct interpretation. The data in Murasugi and Kawamura (2004) indicated that those who had acquired regular scrambling and the property of *zibun* in canonical sentences had no difficulty in interpreting *zibun* in scrambled sentences. The researchers suggested that the reconstruction property of scrambling is acquired around three or four years of age.

To summarize the findings of the L1 studies, it appears that the grammatical knowledge required to process scrambling is acquired quite early, perhaps as early as canonical sentences are acquired. However, Japanese children develop and rely on perceptual strategies to process sentences, just as L1 English children do, and thus their performance in comprehending scrambled sentences remains worse for a certain period of time.[13] While the primary perceptual strategy for L1 English children is the word order strategy, the Japanese perceptual strategy seems to include word-order and case-marker cues.

2.3.2 Processing of OSV sentences by L2 learners of Japanese

Most of the studies that examine the L2 processing of Japanese scrambled sentences have been motivated by the competition model (MacWhinney & Bates, 1989). The competition model assumes that the meaning of a sentence is interpreted in real time by integrating the information provided by different linguistic cues such as word order, morphology (e.g., agreement, case markers), and lexical semantics (e.g., animacy). The cues *converge* when they simultaneously assign the same thematic relation whereas they *compete* when they designate a different thematic relation. For instance, in the sentence, *Two cats are chasing a ball*, the cues (i.e., word order, agreement, and animacy) converge as they simultaneously assign the agent role to the NP *two cats*. On the other hand, in *A ball is chasing two cats*, the cues compete as word order and agreement disagree with animacy in the theta role assignment. According to the model, the relative weights of the cues differ among languages, and thus L2 acquisition entails learning the cue weights of the target language. Since grammatical relations of the NPs in Japanese are ultimately determined by case markers whereas in English they are mostly determined by word order, the difference between the two languages has provided opportunities for researchers to empirically test the competition model.

Kilborn and Ito (1989) investigated the use of word order and case-marking cues in sentence comprehension by novice and advanced L1 English learners of Japanese and by the native

[13] Mizumoto (2009) proposes that the lower comprehension performance of scrambled sentences can be attributed to developing working memory capacity.

speaker controls. The results showed that native speakers relied on the case marker *–ga*, when it was available, to identify the agent in a word sequence, while mostly ignoring the word order information. Advanced L2 learners showed a similar pattern, although less consistently than the native speakers. Novice L2 learners, on the other hand, relied mostly on the word order cue in the identification of the agent in a word sequence, ignoring the case markers. Based on the data, Kilborn and Ito suggested that the novice L2 learners seemed to take “a short cut to sentence interpretation” (p. 284), using the SOV template.

Sasaki (1994) examined the use of word order, case marker, and animacy cues among three groups: beginning and intermediate L1 English learners of Japanese and advanced L1 Japanese learners of English. Sasaki also observed “meta-transfer” as proposed in Kilborn and Ito (1989), that is, the overuse of the SOV template by L1 English learners of Japanese although SOV itself is not the canonical word order in English. Contrary to Kilborn and Ito, however, Sasaki observed that the overuse of the SOV template was more prominent among intermediate learners than beginning learners.

Rounds and Kanagy (1998), also from the perspective of the competition model, investigated the influence of word order and case-marking particles on the comprehension of NNV sequences in Japanese by L1 English children learning Japanese as an L2 in an immersion context. The children were instructed to listen to tape-recorded NNV strings, and to choose the picture that best described the string. The researchers found that, for non-case-marked strings, children preferred to choose the first noun of the string as the agent (76% of the time, overall) but that such a “first-noun-as-agent” preference seemed to have been reinforced as children gained more exposure to Japanese. While the kindergarten and first grade children chose the first noun as the agent only 60% of the time, older children chose the first noun as the agent at a higher rate. It was also found that, even when scrambled sentences were clearly case-marked, the children still tended to choose the first noun as the agent (89.8 % of the time). This tendency was also reinforced as the children's exposure to Japanese increased. While the kindergarten and first grade children chose the *–o* case-marked first noun as agent 74.6% of the time, the sixth and seventh grade children chose it as agent 95.1% of the time. Thus, while the advanced Japanese learners in Kilborn and Ito (1989) learned to use case markers for the assignment of the agent, the children in Rounds and Kanagy (1998) did not show development in the same direction.

The participants in the studies reviewed above were all native speakers of English. Therefore, it is not clear whether their tendency to apply the SOV template in comprehending L2 Japanese is the result of the transfer of the L1 strategy or of a general processing strategy shaped by the typical feature of Japanese (i.e., SOV). Koda (1993) attempted to clarify the point by comparing the comprehension of SOV/OSV sentences by L2 Japanese learners of three different L1s, Chinese, English, and Korean. Chinese, like English, is a language that depends heavily on word order for case assignment. Korean, on the other hand, has postpositional case markers, similar to Japanese, and thus allows scrambling. Koda tested the comprehension of four types of sentences (canonical with and without particles, scrambled with and without particles) to examine if L1 cognitive strategies are transferred in L2 text comprehension. The results showed that all the L1 groups performed better when case-marking particles were present. However, while L1 Chinese and L1 English participants were severely impaired by the OSV order in their comprehension, L1 Korean participants performed equally well for both canonical and scrambled sentences. Given the result, the dependence on the "NP-*ga* NP-*o* V" template of L1 English learners of Japanese may be attributed to the transfer of L1 cognitive strategies.[14]

[14] Iwasaki (2003) proposes that Koda's (1993) results can be interpreted as positive transfer of L1 processing strategy by L1 Korean learners rather than negative transfer of L1 Chinese and English processing strategy.

To the best of the author's knowledge, there have been very few studies conducted on the production of scrambled sentences by L2 learners. Doi and Yoshioka (1987, 1990; cited in Kawaguchi, 2005 and Iwasaki, 2003) investigated the acquisition order of the topic marker –*wa*, the nominative marker –*ga*, and the accusative marker –*o* among L2 leaners of Japanese by calculating the accuracy rate of the three particles in their speech. Their results indicated that –*wa* was acquired before –*ga* and –*o*, and that –*o* was acquired before –*ga*. They also suggested that L2 learners of Japanese become able to produce sentences in free word order as soon as they acquire –*ga* and –*o*, but according to Iwasaki (2003), they did not directly test this point.

So far, the most thorough examination of L2 learners' production of scrambled sentences is provided in Iwasaki (2000, 2003). Not only did Iwasaki investigate production of OSV sentences, she also included different tasks to separately assess L2 learners' grammatical knowledge and production performance on scrambling. Iwasaki used a picture description task to assess L2 learners' production of OSV sentences in real time, and a fill-in-the-blank task and a grammaticality judgment task to assess the learners' grammatical knowledge of particles as a basis to process scrambled sentences. The participants in Iwasaki's study were 31 L1 English learners of Japanese who were divided into three different proficiency groups.

In the picture description task, the participants of all levels were similarly less accurate in producing scrambled sentences. Iwasaki observed a number of instances in which the participants made errors such as "O-*ga* S-*o* V" and "O-*wa* S-*o* V", indicating that the L2 learners used the "NP1-*ga* NP2-*o*" template to produce sentences. The results of the fill-in-the-blank task revealed that the L2 Japanese learners were less accurate in filling out the particles for OSV sentences, regardless of their proficiency. Iwasaki also observed rather large individual differences in the participants' knowledge of scrambling; some of the participants did not seem to know that scrambling is permissible in Japanese. The result of the grammaticality judgment task indicated that the L2 Japanese learners made more errors and took longer to judge OSV sentences than SOV sentences. There was no effect of proficiency in the grammaticality judgment task, either. Since the comparison of the picture description task and the fill-in-the-blank task revealed that there were instances where participants clearly demonstrated their knowledge of the OSV structure yet made errors that were based on the SOV template in producing scrambled sentences, Iwasaki concluded that some production errors of OSV sentences can be attributed to the processing strategy.

To summarize, comprehension studies of L2 Japanese scrambling have found that L2 learners do rely on the SOV template to comprehend simple transitive sentences just as do L1 Japanese children. A crosslinguistic study by Koda (1993) suggested that dependence on the SOV template may be attributed to learners' L1 processing strategy. However, because most of the comprehension studies have been conducted within the framework of the competition model, and therefore they have been mostly concerned about the investigation of relative cue strengths between L1 and L2, they have not addressed L2 learners' grammatical knowledge of scrambling. While there have been several L2 comprehension studies on Japanese scrambling, there have been very few studies that examined L2 production of scrambled sentences. Iwasaki (2000, 2003) is notable in that it examined the gap between L2 learners' knowledge and performance during production of OSV sentences by setting up different tasks. As for the relation between general proficiency in Japanese and processing of scrambling, some studies have found a positive relation between the two whereas others have found persistent processing difficulty with scrambling at a higher proficiency level.

2.4 Rationale and purposes of the present study

The present study was aimed at examining the relation between L2 Japanese learners' grammatical knowledge and production performance, thereby bridging the gap in the previous literature on L2 processing of scrambled sentences in Japanese. In order to do so, fill-in-the-blank and picture description tasks were adopted from Iwasaki (2000, 2003). Task 1 (fill-in-the-blank) aimed at assessing L2 learners' grammatical knowledge of nominative and accusative case markers, the basis for processing scrambling of simple transitive sentences. Task 2 (picture description) attempted to assess L2 learners' performance during their production of scrambled sentences. The rationale for the two tasks is the same as Iwasaki's. In Task 1, sentence items were printed on paper and the participants were freely able to go back to previous items if they needed to. Therefore, it was assumed that the participants were fully able to make use of their grammatical competence in completing the task. On the other hand, Task 2 was timed, and the participants needed to describe the pictures as they were presented on a computer monitor without preparation. It was thus assumed that the task was able to assess the participants' performance during their production of OSV sentences.

The focus of the present study was simple transitive sentences whose verbs take accusative-marked NPs as their internal arguments. On the other hand, Iwasaki's (2000, 2003) objectives were more extensive (i.e., production of postpositional particles in seven different types of verbs), and the accusative-marked transitive sentence was just a subset in her materials. Thus, although the two tasks were adopted from Iwasaki's, changes were made to accommodate the difference and to examine L2 learners' performance on simple transitive sentences more fully. Key changes are the following:

a) Inclusion of reversible sentences. Iwasaki grouped her participants in three proficiency levels and gave a different set of items to each proficiency group for the purpose of accommodating the participants' general proficiency in Japanese. While her materials included reversible sentences, the simple transitive sentences given to all the participants were non-reversible. In order to examine the effects of reversibility on the participants' performance on the tasks, reversible sentences were also given to all the participants in the present study.

b) Inclusion of a larger number of participants. The 31 NNS participants in Iwasaki's study were grouped into Low, Mid, and High groups which included 15, 10, and 6 participants, respectively. Iwasaki did not observe the effects of proficiency on L2 learners' accuracy in processing OSV sentences and attributed the absence of the proficiency effects to rather large individual differences. For the purpose of investigating the relations between L2 learners' general proficiency in Japanese and their processing of the OSV sentences, the present study attempted to recruit a larger number of NNS participants.

c) Order of the tasks. In Iwasaki's study, the picture description task was given to the participants before the fill-in-the-blank task, whereas the fill-in-the-blank task was administered first in the present study. The change of task order was due to the concern that some participants might not even attempt to produce OSV sentences if the picture description task was administered first, given that the scrambling structure is generally not taught or practiced explicitly in classroom instruction. It was felt that, by providing the fill-in-the-blank task first, the participants would be able to activate their knowledge of scrambling more fully before they worked on the picture description task, making the comparison of the two tasks more pertinent. The change in task order also seems to have helped limit the variation in the participants' responses. (As will be discussed later, the NS participants were not given the fill-in-the-blank task, and they used non-target structures quite frequently to describe the pictures.)

d) Scoring. Iwasaki's scoring was based on each positional particle. As each sentence item had two NPs, each item provided two scoring points. While it was an optimal scoring scheme for the objective of Iwasaki's study (i.e., the production of particles), it may be less optimal for the purpose of evaluating performance on the scrambling structure. This is because a response such as "NP-*ga* NP-*ga* V" (in response to a scrambling stimulus) would automatically receive a half point using the scoring based on each particle although this response does not prove an understanding of the scrambling structure. For this reason, each sentence item was scored as a whole in the present study.

The specific research questions addressed in the present study are the following:

a) Do L2 learners have the grammatical knowledge to process scrambling?
b) Do L2 learners experience difficulty in producing OSV sentences as compared to SOV sentences?
c) If L2 learners experience difficulty in producing OSV sentences, is the source of difficulty the lack of grammatical knowledge or a processing problem?
d) If there is a gap between L2 learners' grammatical knowledge and production performance, can it be attributed the NNV strategy?
e) What is the relationship between the grammatical knowledge/production of the scrambling structure and L2 learners' general proficiency in Japanese? Are more proficient learners better at processing OSV sentences?

f) What are the effects of reversibility on L2 learners' processing of OSV sentences? Do L2 learners integrate the animacy contrast during their processing of the scrambling structure?

2.5 Method for Task 1

Task 1 (fill-in-the-blank task) was conducted for the purpose of assessing L2 learner's grammatical knowledge of scrambling.

2.5.1 Participants

The fifty-four NNS participants described in Chapter 1 participated in Task 1. Task 1 was not administered to the NS participants, as the task was to assess the knowledge of rather simple grammatical structures, and it was expected that the NS participants would obtain almost perfect scores regardless of the conditions.

2.5.2 Materials and design

There were four types of sentence stimuli as shown below. (See Appendix 2A for a list of target items in Task 1.)

(2.5) a. *Reversible canonical:*
Furansujin-ga Kyoko-o mita.
French person-Nom -Acc saw
'The French person saw Kyoko.'

b. *Reversible scrambled:*
Kyoko-o Fransujin-ga mita.
-Acc French person saw

c. *Non-reversible canonical:*
Onnanohito-ga repooto-o kaita.
woman-Nom report-Acc wrote
'The woman wrote the report.'

d. *Non-reversible scrambled*
Repooto-o onnanohito-ga kaita.
report-Acc woman-Nom wrote

The scrambling of simple transitive sentences such as above is rather infrequent in Japanese. On the other hand, at least intuitively, the scrambling of locative NPs as well as other oblique NPs appears more frequent. To examine whether such a difference in frequency might influence grammatical knowledge among L2 learners, sentences with locative NPs such as below were also included in the item set.

(2.6) a. *Locative canonical:*
Ichiro-ga apaato-de neta.
-Nom apartment-Loc slept
'Ichiro slept in the apartment.'

b. *Locative scrambled:*
Apaato-de Ichiro-ga neta.
apartment-Loc -Nom slept

To prepare the sentence set, 8 verbs that can be used reversibly, 8 verbs that can be used non-reversibly, and 6 verbs that can be used with locative NPs were chosen, and each verb was used twice to make 44 canonical sentences in total (16 reversible, 16 non-reversible, and 12 locative). The pairs of the sentences that shared the same verb were split into two equivalent lists. At this point, the scrambled sentences were created by switching the nominative and accusative-marked NPs in each of the two equivalent lists. The scrambled sentences which were made from the canonical sentences on one list were added to the other list, and vice versa, so that participants would not encounter sentences with the same NPs-verb combination during the task. Thirty-six filler sentences with various structures were added to each item list. Thus, each item list consisted of 8 sentences in each of the four sentence types in (2.5) above, 6 each of locative canonical/scrambled sentences, and 36 filler sentences, for a total of 80 items. The items in each list were pseudo-randomized so that sentences of the same type would not be adjacent to each other. In an effort to alleviate the influence of the item presentation order, two pseudo-randomized versions were created for each item list.

All the verbs, inanimate nouns and location nouns used in the sentence items were chosen from *Nakama 1 & 2 2nd Edition* (Hatasa, Hatasa, & Makino, 2009), the textbooks used in the beginning and intermediate Japanese courses at the University of Arizona. About half of the animate nouns were also chosen from the textbooks, but the other half were common Japanese names for the purpose of minimizing the repetition of the same nouns in the item set. One of the nouns in each reversible sentence was a regular noun (such as ‘girl’ and ‘roommate’) while the other noun was a name. In the effort to minimize the effect of this regular noun/name contrast on the participants’ performance, regular nouns were nominative-marked in half of the reversible sentences and names were nominative-marked in the other half. Half of the non-reversible sentences had regular nouns as their subjects, and the other half had names as subjects. The subjects of the locative sentences were all names.

The item lists were presented in a paper-and-pencil format. Each item list was printed on 10 sheets of paper; the first page was the instruction, and the sentence items were printed on the rest. Below is an example of a trial. (The example below is a non-reversible canonical sentence. The underlined elements were included for demonstration purposes only and were not included in the actual materials.)

(2.7)

The woman wrote the report. レポート（ <u>を</u> ）女の人（ <u>が</u> ）書いた。 *report* *woman* *wrote* <u>Repooto-(o) onnanohito-(ga) kaita.</u> <u>report-(Acc) woman-(Nom) wrote</u>

The sentence items were written in Japanese orthography with *furigana* (*hiragana* as phonetic guide) over *kanji*. All the postpositional particles in the sentence items were replaced with parentheses for the participants to fill in based on the English sentence above. Glosses were provided under each word in a sentence item for the purposes of providing the meaning of each word as well as making the order of the nouns more salient for the participants.

2.5.3 Procedure

The participants were instructed to read the English sentence first and to fill in the parentheses with Japanese particles so that the completed Japanese sentence would have the same meaning as the English sentence. While it was written in the instructions, the researcher also asked the participants to check that the Japanese sentence has the same meaning as the English sentence after they filled in the parentheses. The possible choices of the particles were listed on the instructions. The use of the topic marker –*wa* was discouraged for this task.

2.5.4 Scoring and analysis

Each item was scored for its accuracy. Responses were given the score of 1 if they were correct, and 0 if they were incorrect. Responses were judged correct only if both of the particles in a given sentence were correct. Therefore, the correct combination of particles were *ga-o* for the canonical transitive sentences, *o-ga* for the scrambled transitive sentences, *ga-de* for the canonical locative sentences, and *de-ga* for the scrambled locative sentences. As will be observed, however, the NNS participants frequently used the *–ni* particle to mark the direct object of the transitive sentences instead of the correct *–o* particle, especially for the reversible sentences. In such instances, a participant would fill in *ga-ni* for the two parentheses of the canonical transitive sentences and *ni-ga* for the scrambled counterparts, and the aforementioned scoring scheme would give the scores of 0 for both of the responses while these responses clearly demonstrate a grammatical knowledge of the scrambling construction. To alleviate this problem, another set of scoring was conducted which included the *–ni* marking of the direct object as a correct answer.[15] Because the marking of the location nouns with *–ni* (instead of correct *–de*) was also frequent, such responses were also considered correct for the locative sentences in this scoring.[16] In this second coding, the use of a particle other than *–o* or *–ni* to mark the accusative NP was also considered correct if the same particle was used to mark the accusative NPs of the canonical and scrambled sentences that shared the same verb. For the ease of explanation, the latter scoring will be called the "lax" scoring and the former "strict" scoring hereafter. The results based on the "lax" scoring will be reported first in the results section below as this scoring seems to better reflect the participants' knowledge of the scrambling construction than the "strict" scoring does.

In addition to the scoring above, raw frequencies of the actual particles that each participant used to complete the scrambled sentences were counted. This was for the purpose of observing

---

[15] The *–ni* particle has several functions, and is used transitively in some instances. For example, verbs such as *au* ('to meet') and *kisu-suru* ('to kiss') take *–ni* marked NPs as object. Sugimura (2002) maintains that the prototypical meaning of *–ni* is "goal". From the cognitive linguistic perspective, Moriyama (2008) suggests that, while the *–o* marked accusative NP is under the control of the action caused by the nominative NP and thus is the "passive participant", the *–ni* marked dative NP is the "active participant" of the action.

[16] Both *–ni* and *–de* particles mark location NPs. While *–ni* signifies the goal location, *–de* signifies the location where the action takes place. Thus, for actions such as 'go to + location' and 'enter + location', the location NP is marked with *–ni*, whereas *–de* is used to mark the location NP for such actions as 'play + location' and 'study + location'.

response patterns on the individual level. The frequencies of "NP1-*ga*" and "NP2-*o/ni*" in the items of the scrambling condition were also counted for the purpose of observing whether there was tendency to erroneously add –*ga* after the first NPs and/or –*o* or –*ni* after the second NPs.

The accuracy data were analyzed using linear mixed-effects modeling in R (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000). Unlike the F1.F2 analysis, this method allows the analysis of the data from each trial without aggregating over subjects or items. Because locative sentences were included in Task 1 only, the transitive and locative sentences were analyzed separately. The accuracy data for transitive sentences were analyzed first by proficiency group, using Reversibility and Word order as fixed effects. For the purpose of analyzing the differences among the proficiency groups, a subsequent analysis was conducted with the addition of Proficiency as a fixed effect. For the analysis of the locative sentences, Word order and Proficiency were entered as fixed effects. For all the analyses, Subjects and Items were random effects.

For this and all the subsequent analyses using the mixed-effects modeling, the models which included interactions of fixed effects were compared to the model without interaction using a likelihood ratio test, and the model that best fit the data was chosen. This model was then compared with models that included by-subject and by-item random slopes for the effect of Reversibility and/or Word order, using a likelihood ratio test.[17] The analyses were conducted using both models with and without random slopes. The results using random slopes models will be reported if they are substantially different from those without random slopes. The models used for each analysis will be provided. P-values were obtained using the *summary* function of the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2013). If the model used to report

[17] The random slopes for Reversibility were included only for Subject since the items were separate for the reversible and non-reversible conditions.

the results included interactions of fixed effects, the p-values for main effects (for a factor with two levels) and/or simple effects (for a factor with three levels) were obtained through the model without the interactions using *lmerTest*. If the model used to report the result did not include interactions, p-values for interactions were obtained by comparing it with the model that included the interactions using likelihood ratio tests.

2.6 Results of Task 1

2.6.1 Mean response accuracy data based on the "lax" scoring

2.6.1.1 Transitive sentences

Figure 2.1 presents the mean accuracy rates for the transitive sentences by the three proficiency groups. The mean accuracy rates here are based on the "lax" scoring.

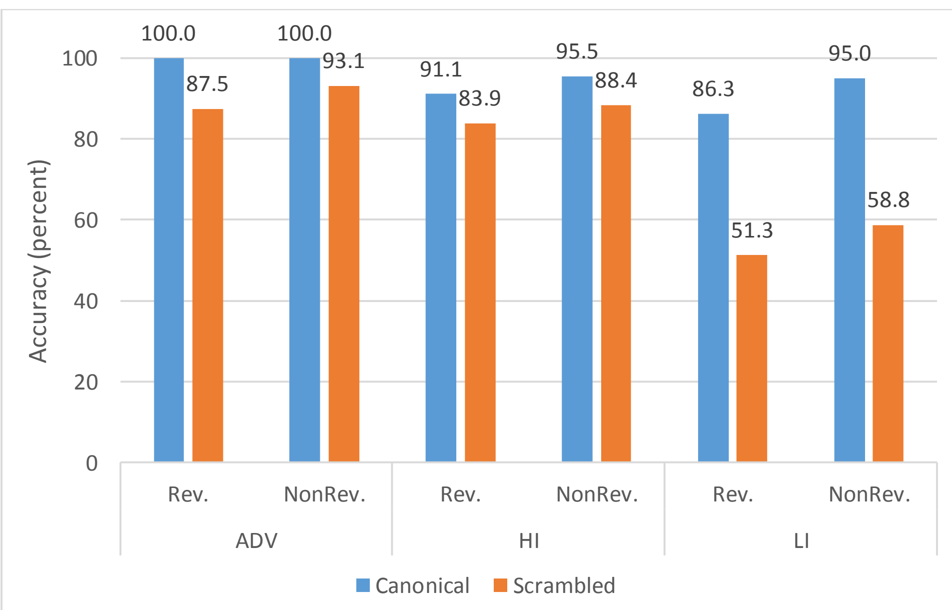


Figure 2.1: Mean accuracy rates for the transitive sentences based on the "lax" scoring in Task 1

The analysis of each proficiency group indicated no significant interaction of Reversibility and Word order although it approached significance for the ADV group ($\chi^2(1) = 3.4356$, $p = 0.0638$). Thus, the following model was used to analyze the data of each group: `Accuracy ~ Reversibility + Word order + (1 | subject) + (1 | item).`

For the ADV group, the main effect of Word order was significant ($t = 6.395$, $p < 0.0001$), and the main effect of Reversibility approached significance ($t = 1.848$, $p = 0.0651$). On the other hand, when the dataset was analyzed using the model with by-subject and by-item random slopes for the effect of Word order, there was a large decrease in the t value of the Word order effect ($t = 2.281$, $p = 0.0335$).[18] There was also a slight decrease of the Reversibility effect with the random slopes model ($t = 1.322$, $p = 0.1897$).

For the HI group, the main effect of Word order was significant ($t = 2.761$, $p = 0.006$), and the main effect of Reversibility approached significance ($t = 1.725$, $p = 0.0852$). However, with an analysis using a random slopes model (by-subject for the effect of Word order), the main effect of Word order turned out non-significant ($t = 1.364$, $p = 0.1957$).

For the LI group, the main effects of Word order as well as Reversibility were significant (Word order: $t = 13.561$, $p < 0.0001$; Reversibility: $t = 2.874$, $p = 0.0074$). However, again, an analysis with a random slopes model (by-subject for the effect of Word order) indicated a substantial decrease of the effect of Word order although it was still highly significant ($t = 4.563$, $p = 0.0002$).

For the purpose of observing the differences between the proficiency groups, an analysis that included Proficiency was also carried out. The following model with random slopes was used: `Accuracy ~ Reversibility + Word order * Proficiency + (1 +`

[18] The difference between the “by-subject and by-item” random slopes model and the “by-subject only” random slopes model was only marginally significant ($\chi^2(2) = 6.1561$, $p = 0.0461$).

`Reversibility + Word order | subject) + (1 | item)`. The analysis indicated that the main effects of Reversibility and Word order were significant (Reversibility: $t = 2.802$, $p = 0.0082$; Word order: $t = 4.723$, $p < 0.0001$). The effect of Proficiency was marginally significant between the ADV and LI groups ($t = 2.05$, $p = 0.0456$), but it was not significant either between the ADV and HI groups or the HI and LI groups (ADV-HI: $t = 1.73$, $p = 0.0898$; HI-LI: $t = 0.131$, $p = 0.8963$). As the model indicates, there was also a significant interaction between Word order and Proficiency ($\chi^2(2) = 13.518$, $p = 0.0012$).

For the canonical sentences, the mean accuracy rate of the ADV group was significantly higher than that of the LI group ($t = 2.339$, $p = 0.0234$) but not significantly higher than those of the HI group ($t = 1.704$, $p = 0.0945$). There was no significant difference in the mean accuracy rates of the canonical sentences between the HI and LI groups ($t = 0.418$, $p = 0.6775$). For the scrambled sentences, no significant difference was observed between the ADV and HI groups ($t = 0.417$, $p = 0.6787$) but the mean accuracy rates for the scrambled sentences were significantly different between the ADV and LI groups as well as between the HI and LI groups (ADV-LI: $t = 3.368$, $p = 0.0014$; HI-LI: $t = 3.274$, $p = 0.0019$). Finally, the interaction of Word order and Proficiency was significant between the ADV and LI groups as well as between the HI and LI (ADV-LI: $t = 3.368$, $p = 0.0014$; HI-LI: $t = 3.274$, $p = 0.0019$), where the difference between the canonical and scrambled sentences were significantly larger for the LI group. The interaction was not significant between the ADV and HI groups ($t = 0.217$, $p = 0.8287$). One difference between the previous within-group analyses and this between-group analysis is that the effect of Word order which was significant for the ADV group in the previous analysis (even with the random slopes model) did not reach a level of significance in this three-factor analysis ($t = 1.515$, $p = 0.1359$).

All in all, the results indicated that the participants across proficiency groups were less accurate in their response with scrambled sentences. However, the participants in the LI group were significantly less accurate with the scrambled sentences as compared to the other two groups. As for the effect of reversibility, the reversible sentences were responded to less accurately than the non-reversible sentences. However, the sentence items in the reversible and non-reversible conditions were strictly separate, and therefore, this difference in accuracy rates might be attributed not only to reversibility but to the difference in the items.

2.6.1.2 Locative sentences

Figure 2.2 shows the mean accuracy rates for the locative sentences based on the "lax" scoring. The following model was used to analyze the data: `Accuracy ~ Word order * Proficiency + (1 | subject) + (1 | item)`. As the model suggests, the interaction of Word order and Proficiency was significant ($\chi^2(2) = 23.634$, $p < 0.0001$).

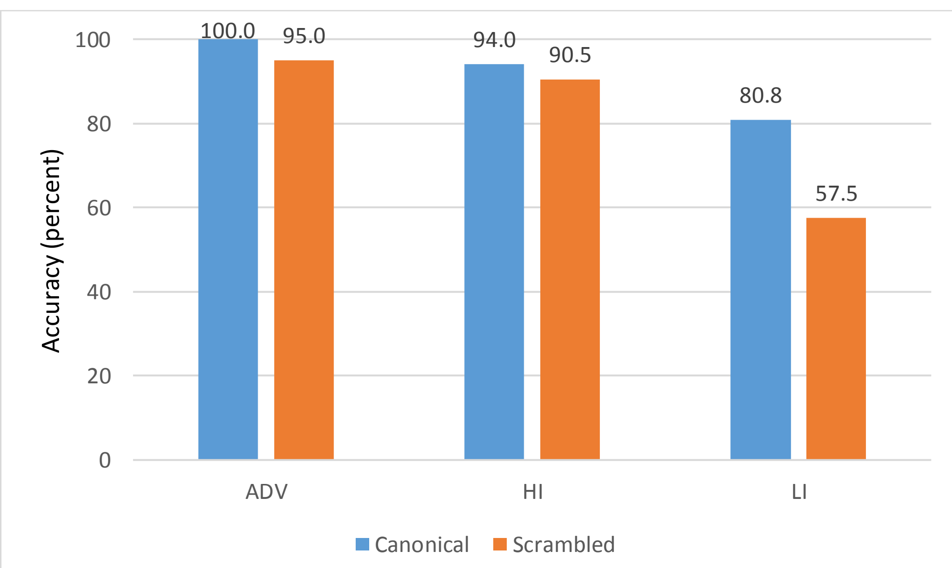


Figure 2.2: Mean accuracy rates for the locative sentences based on the "lax" scoring in Task 1

The effect of Word order was significant with the LI group (t = 7.568, p < 0.0001), but not significant with the ADV and HI groups (ADV: t = 1.628, p = 0.104; HI: t = 0.973, p = 0.331). On the other hand, an analysis using a random slopes model (by-subject for the effect of Word order) indicated a large decrease in the effect of Word order for the LI group (t = 3.086, p = 0.0033). The interaction of Word order and Proficiency was significant between the ADV and LI groups as well as between the HI and LI groups (ADV-LI: t = 4.221, p < 0.0001; HI-LI: t = 4.129, p < 0.0001) but it was not significant between the ADV and HI groups (t = 0.298, p = 0.7654).

The mean accuracy rates of the transitive and locative sentences seem to indicate a similar pattern in that the LI participants were significantly less accurate in completing sentences in the scrambled word order.

2.6.2 Mean response accuracy data based on the "strict" scoring

This sections will present the results based on the "strict" scoring, in which the only correct combination of particles were *ga-o* for the canonical transitive sentences, *o-ga* for the scrambled transitive sentences, *ga-de* for the canonical locative sentences, and *de-ga* for the scrambled locative sentences.

2.6.2.1 Transitive sentences

Figure 2.3 shows the mean accuracy rates for the transitive sentences based on the "strict" scoring.

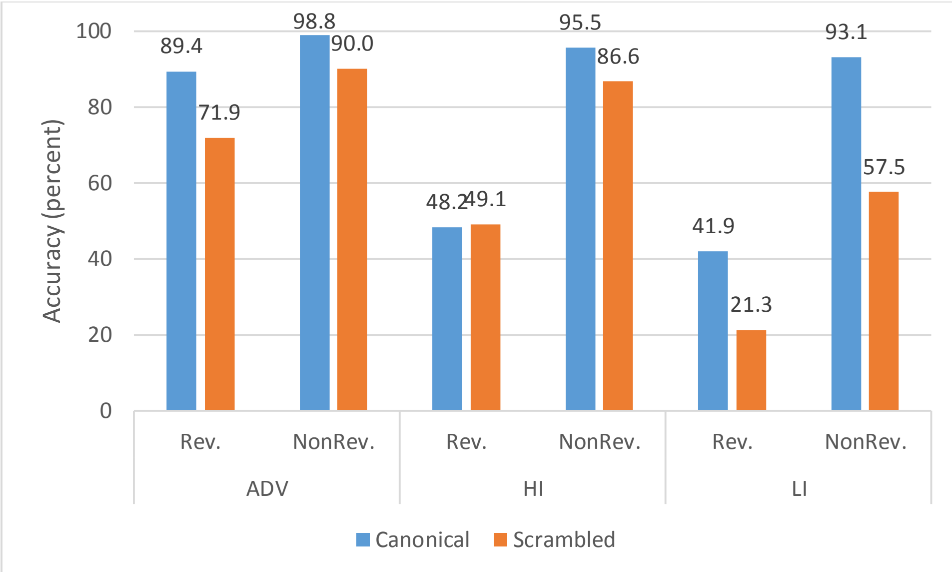


Figure 2.3: Mean accuracy rates for the transitive sentences based on the “strict” scoring in Task 1

For the data based on the “strict” scoring, different linear mixed-effects models were used to analyze the mean accuracy rates of the three proficiency groups. For the analysis of the ADV group data, the following model was used: `StrictAccuracy ~ Reversibility * Word order + (1 | subject) + (1 | item)`. The main effects of the two fixed effects were obtained with the following model: `StrictAccuracy ~ Reversibility + Word order + (1 | subject) + (1 | item)`. The main effects of the two factors were both significant (Reversibility: $t = 3.163$, $p = 0.0036$; Word order: $t = 6.21$, $p < 0.0001$). As the model suggests, the interaction of Reversibility and Word order was significant ($t = 2.076$, $p = 0.0383$). For both the reversible and non-reversible sentences, the effect of Word order was significant (Reversible: $t = 5.872$, $p < 0.0001$; Non-reversible: $t = 2.936$, $p = 0.0035$). On the other hand, when the same dataset was analyzed with a random slopes model (by-subject for the effects of Reversibility and Word order AND by-item for the effect of Word order), the interaction of the two fixed effects did not reach significance ($\chi^2(1) = 2.4811$, $p = 0.1152$). The analysis using a random slopes model without the interaction of the two fixed effects indicated that the main effect of Reversibility was significant (t = 2.696, p = 0.0106) and the main effect of Word order approached significance (t = 2.006, p = 0.0557).

For the HI group, the following model was used: `StrictAccuracy ~ Reversibility + Word order + (1 | subject) + (1 | item)`. The main effect of Reversibility was highly significant (t = 6.387, p < 0.0001), but the main effect of Word order was not significant (t = 1.260, p = 0.208). The interaction of the two factors was not significant ($\chi^2(1) = 2.385$, p = 0.1225). An analysis using a random slopes model did not alter the results for the HI group.

For the LI group, the interaction of Reversibility and Word order was observed, and hence the following model was used: `StrictAccuracy ~ Reversibility * Word order + (1 | subject) + (1 | item)`. The main effects of the two fixed effects were obtained with the following model: `StrictAccuracy ~ Reversibility + Word order + (1 | subject) + (1 | item)`. The main effects of the two factors were both significant (Reversibility: t = 9.673, p < 0.0001; Word order: t = 9.525, p < 0.0001). The interaction was also significant (t = 2.552, p = 0.011), where the difference in the mean accuracy rates between the canonical and scrambled sentences was larger for the non-reversible sentences than the reversible sentences. The simple effect of Word order was significant for the reversible sentences (t = 4.962, p < 0.0001) as well as for the non-reversible sentences (t = 8.571, p < 0.0001). An analysis using a random slopes model (by-subject for the effects of Reversibility and Word order) indicated a large decrease in the simple effect of Word order for both the reversible and non-reversible sentences (Reversible: t = 2.786, p = 0.0103; Non-reversible: t = 4.813, p < 0.0001).

In order to compare the differences among the three proficiency groups, an analysis which included Proficiency as the third fixed effect was carried out. The following model with random slopes was used for the analysis: `StrictAccuracy ~ Reversibility * Word order * Proficiency + (1 + Reversibility + Word order | subject) + (1 | item)`. The main effect of Reversibility was significant ($t = 5.892$, $p < 0.0001$) as was the main effect of Word order ($t = 4.174$, $p = 0.0001$). The effect of Proficiency was significant between the ADV and HI groups ($t = 3.536$, $p = 0.0009$) as well as between the ADV and LI groups ($t = 4.896$, $p < 0.0001$), but it was not significant between the HI and LI groups ($t = 0.907$, $p = 0.3685$). There were significant interactions of Proficiency and Reversibility ($\chi^2(2) = 21.784$, $p < 0.0001$), and of Proficiency and Word order ($\chi^2(2) = 10.087$, $p < 0.0065$), but the interaction of Reversibility and Word order only approached significance ($\chi^2(1) = 3.0764$, $p = 0.0794$). There was also a three-way interaction among the fixed effects ($\chi^2(2) = 13.11$, $p = 0.0014$).

Moving on to the pairwise comparisons of the interactions among the proficiency groups, the interaction of Proficiency and Word order for the reversible sentences was marginally significant between the HI and LI groups ($t = 2.106$, $p = 0.0391$). It approached significance between the ADV and HI groups ($t = 1.8$, $p = 0.0765$), and was not significant between the ADV and LI groups ($t = 0.337$, $p = 0.7372$). For the non-reversible sentences, the interaction between Proficiency and Word order was significant between the ADV and LI groups ($t = 2.898$, $p = 0.0051$) as well as between the HI and LI group ($t = 2.613$, $p = 0.0111$) but not significant between the ADV and HI groups ($t = 0.017$, $p = 0.9861$).

The interaction of Proficiency and Reversibility for the canonical sentences was significant between the ADV and HI groups ($t = 4.356$, $p < 0.0001$) as well as between the ADV and LI groups ($t = 5.298$, $p < 0.0001$), but not between the HI and LI groups ($t = 0.451$, $p = 0.6533$). For the scrambled sentences, the interaction of Proficiency and Reversibility was significant between the ADV and HI groups ($t = 2.242$, $p = 0.028$) as well as between the ADV and LI groups ($t = 2.311$, $p = 0.0237$), but again not significant between the HI and LI groups ($t = 0.145$, $p = 0.8854$).

Finally, the three-way interaction of the fixed effects was significant between the ADV and HI groups ($t = 2.612$, $p = 0.0091$) and between the ADV and LI groups ($t = 3.681$, $p = 0.0002$), but not between the HI and LI groups ($t = 0.728$, $p = 0.4665$).

In comparison to the previous analyses of the "lax" data, the current analyses of the "strict" data indicated that the participants were generally quite accurate in their marking of the accusative NPs in the non-reversible sentences. Indeed, the interaction of Proficiency and Word order for the non-reversible sentences paralleled with the previous analyses, in which the LI group was significantly less accurate with the scrambled sentences as compared to the other two groups. On the other hand, the interactions of Proficiency and Reversibility in the current analyses indicated that the participants in the HI and LI groups were significantly less accurate with the case markers of the reversible sentences. Of particular interest is the HI group, which was comparable with the ADV group in their knowledge of the scrambling construction in the earlier analysis. The same group, however, was more similar to the LI group in their accuracy in the marking of the accusative NPs of the reversible sentences.

2.6.2.2 Locative sentences

The mean accuracy rates for the locative sentences based on the "strict" scoring is shown in Figure 2.4.

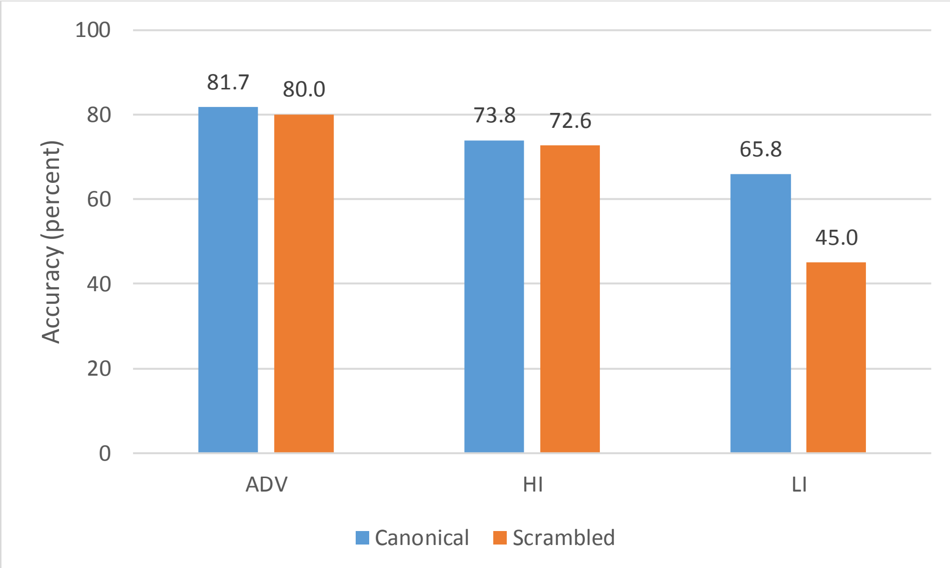


Figure 2.4: Mean accuracy rates for the locative sentences based on the “strict” scoring in Task 1

The following model was used to analyze the data: `StrictAccuracy ~ Word order * Proficiency + (1 | subject) + (1 | item)`. The main effect of Word order was significant (t =3.21, p = 0.0014). The effect of Proficiency was significant between the ADV and LI groups (t = 2.712, p = 0.0091). It approached significance between the HI and LI groups (t = 1.723, p = 0.0909), and it was not significant between the ADV and HI groups (t = 0.738, p = 0.4641). The interaction of Word order and Proficiency was significant ($\chi^2(2)$ = 12.215, p = 0.0022). The simple effect of Word order was significant only for the LI group (t = 4.769, p < 0.0001).

The interaction of Word order and Proficiency was significant between the ADV and LI groups (t = 3.091, p = 0.0021) as well as between the HI and LI groups (t = 2.875, p = 0.0042), but it was not significant between the ADV and HI groups (t = 0.07, p = 0.9444). The result therefore indicated that the LI group was significantly less accurate with the locative scrambled sentences as compared with the other two groups. The results correspond with the previous analysis of the “lax” data.

2.6.3 Raw frequency data

This section will present analyses based on the raw frequencies of the particles that the participants used to complete the sentences in Task 1. The purpose of the analyses is to observe the response patterns on the individual level. As the purpose of Task 1 is to assess L2 learner's grammatical knowledge on scrambling, the observation here will focus mostly on the items in the scrambling condition.

2.6.3.1 Reversible sentences

Table 2.1 presents the raw frequencies of the particles that the participants in each group used to complete the reversible scrambled sentences. There were eight opportunities for each participant. There were 9 participants in the ADV group, 4 participants in the HI group, and 16 participants in the LI group (a total of 29 participants) who marked the first NP with –*ga* and the second NP with a particle other than –*ga* (i.e., resulting in an SOV form) at least once. As Task 1 was intended to assess the participants' grammatical knowledge, the use of a processing strategy such as the SOV template was not assumed originally. Given the number of participants who erroneously used SOV forms in completing the OSV sentences, however, it is quite likely that the participants applied the SOV template in completing the task. There were 68 instances of errors with an SOV form (i.e., NP1-*ga*, NP2-another particle).

Table 2.1: Raw frequencies of the particles used for the reversible sentences in Task 1

#	Groups	*o-ga*	*(*)ni-ga*	**ga-o*	**ga-ni*	**ni-o*	other	#	Groups	*o-ga*	*(*)ni-ga*	**ga-o*	**ga-ni*	**ni-o*	other
NNS02	ADV	8						NNS30	HI	6	2				
NNS03	ADV	5	1	1	1			NNS31	HI	2	1	1	2	2	
NNS08	ADV	6	2					NNS32	HI	6	1	1			
NNS10	ADV	2	4		1	1		NNS33	HI	5	3				
NNS11	ADV	7	1					NNS37	HI	4				1	*o-ni, de-ni, to-o*
NNS18	ADV		4	1	2	1		NNS38	HI	2	6				
NNS24	ADV	8						NNS43	HI		6			2	
NNS26	ADV	8						NNS04	LI		1		1	4	*kara-o, to-ni*
NNS27	ADV	7					*ga-de*	NNS05	LI			2	5	1	
NNS29	ADV	7	1					NNS06	LI	1	7				
NNS40	ADV	7	1					NNS07	LI	5	2		1		
NNS42	ADV	7	1					NNS13	LI		8				
NNS44	ADV		3	1		3	*o-ni*	NNS15	LI	1	7				
NNS45	ADV	7					*kara-ga*	NNS16	LI	1	6		1		
NNS46	ADV	5		2	1			NNS17	LI		1	1	1	5	
NNS47	ADV	5	2	1				NNS21	LI	4	1	1	1	1	
NNS48	ADV	6	1	1				NNS23	LI	5		2	1		
NNS52	ADV	6	1	1				NNS28	LI	6	1	1			
NNS53	ADV	6	2					NNS34	LI		7				*ga-ga*
NNS54	ADV	8						NNS35	LI	3		2			*ni-karaX2, ga-kara*
NNS01	HI	6	2					NNS36	LI			5	1	2	
NNS09	HI	5	2				*to-ga*	NNS39	LI	2	2	1		3	
NNS12	HI	5	2				*kara-ga*	NNS41	LI			5		1	*no-oX2*
NNS14	HI	6	2					NNS49	LI		1	5		2	
NNS19	HI	2	6					NNS50	LI	5	2		1		
NNS22	HI		4	1	1	2		NNS51	LI	1		4	1	2	
NNS25	HI	6	1	1				NNS55	LI		2	1	2	3	

Note: "NP1-*ni* NP2-*ga*" was considered correct in the "lax" scoring while it was considered incorrect in the "strict" scoring.

Figure 2.5 is a histogram showing how many participants in each proficiency group completed the reversible scrambled sentences correctly (based on the “lax” scoring).

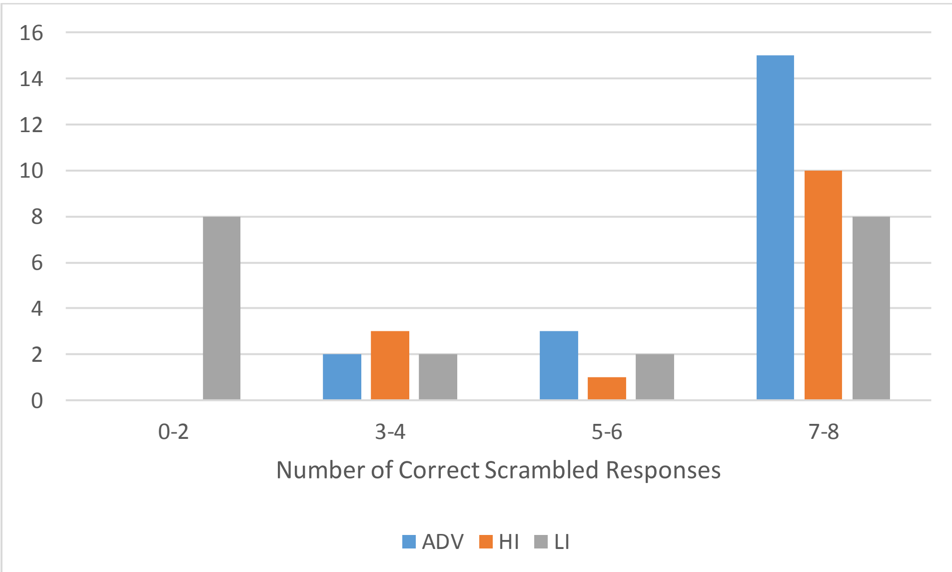


Figure 2.5: The number of participants by proficiency group who completed the reversible scrambled sentences correctly in Task 1

If we define “constant error” as completing the items in an OSV form four times (i.e., 50%) or less, there were 2 participants in ADV, 3 participants in HI, and 10 participants in LI who made constant errors. Looking at the data from a different perspective, although the LI participants were significantly less accurate in completing the scrambled sentences, 10 out of 20 participants in the group completed the reversible scrambled sentences in an OSV form five times or more frequently, suggesting rather large individual differences within proficiency groups. There were three participants in the LI group (NNS05, 36, and 41) who never filled in the parentheses in such a way that the completed sentence had the OSV word order. These participants might not have known that the scrambling structure is permissible in Japanese.

As mentioned previously, many participants used the *–ni* particle, instead of the correct *–o* particle, to mark the accusative NPs in the reversible sentences (cf. Table 2.1). There were 11

ADV participants, 13 HI participants, and 17 LI participants who used a *–ni* particle at least once in marking the accusative NPs in the reversible canonical sentences, indicating that the participants might have been less familiar with the reversible verbs used in the task. The exact reason for the use of the *–ni* particle is not known. However, the *–ni* particle often marks the goal or recipient in a sentence, and because the accusative NPs in the reversible sentences can be interpreted as the recipient of the action denoted by the verb (e.g., see, bully, scold), the participants might have employed *–ni* when they were not sure which particle should be used for a particular reversible verb.

One frequent type of error in completing the reversible scrambled sentences was to mark the first NP with *–ni* and the second NP with *–o*. Again it is not known exactly why this type of error was frequent. Because the first NP in the scrambled sentences can be interpreted as the recipient of the action, the participants might have marked the NP with *–ni*. They then might have marked the second NP with *–o* because this is a preverbal position where the *–o* particle is often used.

While many of the errors in completing the scrambled sentences were caused by the marking of the first NP with *–ga* and the second NP with some other particle (i.e., the SOV form), there were other types or errors, as evidenced in the *ni-o* errors above. For the purpose of examining whether the errors in completing the scrambled sentences were caused by the marking of the first NP with *-ga* or the marking of the second NP with an object marker, the raw frequencies of the cases in which the first NP was marked with a subject marker and the second NP was marked with an object marker were counted. The subject marker was defined as *–ga* and the object markers were *–o* or *–ni* in this analysis. There were 69 instances of wrong NP1-*ga* and 110 cases of wrong NP2-*o*/*ni*. A Chi-square test indicated that the participants supplied *–o* or *–ni* for the second NPs significantly more frequently than *–ga* for the first NP ($\chi^2(1) = 9.391$ $p = 0.002$).

2.6.3.2 Non-reversible sentences

Table 2.2 presents the raw frequencies of the particles that each participant used to complete the non-reversible scrambled sentences. There were eight opportunities for each participant. There were 3 participants in the ADV group, 2 participants in the HI group, and 10 participants in the LI group who marked the first NP with *–ga* and the second NP with another particle at least once, resulting in an SOV form. There were a total of 59 instances of errors with an SOV form.

Table 2.2: Raw frequencies of the particles used for the non-reversible sentences in Task 1

#	Groups	*o-ga*	*(*)ni-ga*	**ga-o*	other
NNS02	ADV	8			
NNS03	ADV	8			
NNS08	ADV	8			
NNS10	ADV	6	1	1	
NNS11	ADV	8			
NNS18	ADV		3	5	
NNS24	ADV	8			
NNS26	ADV	8			
NNS27	ADV	8			
NNS29	ADV	8			
NNS40	ADV	8			
NNS42	ADV	8			
NNS44	ADV	2	1	5	
NNS45	ADV	8			
NNS46	ADV	8			
NNS47	ADV	8			
NNS48	ADV	8			
NNS52	ADV	8			
NNS53	ADV	8			
NNS54	ADV	8			
NNS01	HI	8			
NNS09	HI	8			
NNS12	HI	8			
NNS14	HI	8			
NNS19	HI	8			
NNS22	HI		2	4	*de-o, de-ga*
NNS25	HI	8			

#	Groups	*o-ga*	*(*)ni-ga*	**ga-o*	**ga-ni*	**ni-o*	**o-ni*	other
NNS30	HI	8						
NNS31	HI	8						
NNS32	HI	7		1				
NNS33	HI	8						
NNS37	HI	4					3	*ni-de*
NNS38	HI	8						
NNS43	HI	6				2		
NNS04	LI	1				4		*to-oX2, no-ga*
NNS05	LI			8				
NNS06	LI	8						
NNS07	LI	8						
NNS13	LI	7					1	
NNS15	LI	8						
NNS16	LI	7	1					
NNS17	LI	2			2	1	3	
NNS21	LI	7						*ga-no*
NNS23	LI	7		1				
NNS28	LI	8						
NNS34	LI	8						
NNS35	LI	8						
NNS36	LI			8				
NNS39	LI	4		3				*ga-ga*
NNS41	LI			6				*no-o, ga-no*
NNS49	LI		1	6		1		
NNS50	LI	7						*de-ga*
NNS51	LI	2		2		2	2	
NNS55	LI			3	2	3		

Note: "NP1-*ni* NP2-*ga*" was considered correct in the "lax" scoring while it was considered incorrect in the "strict" scoring.

Figure 2.6 is a histogram showing how many participants in each proficiency group completed the non-reversible OSV sentences correctly (based on the "lax" scoring).

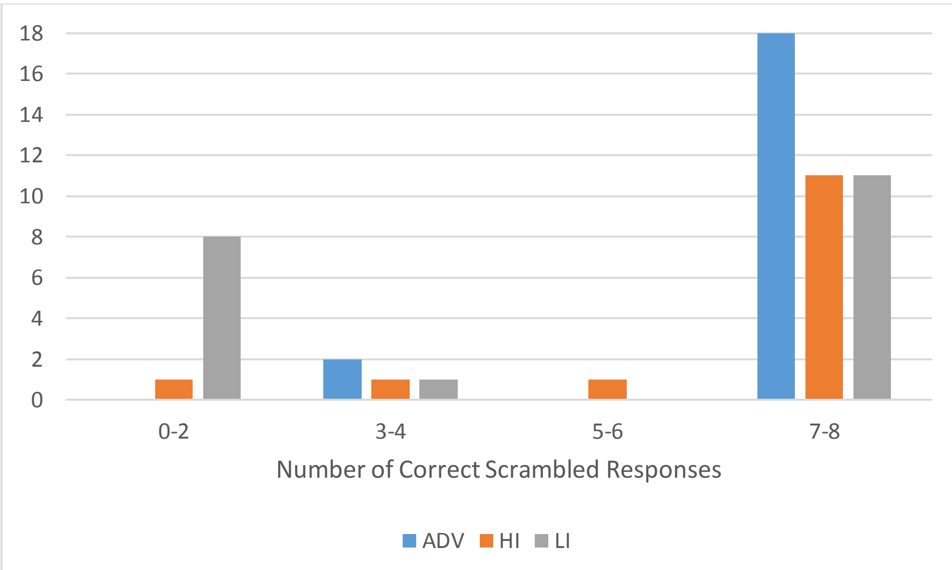


Figure 2.6: The number of participants by proficiency group who completed the non-reversible scrambled sentences correctly in Task 1

If we define "constant error" as completing the items in an OSV form less than four times (i.e., 50%) or fewer, there were 2 participants in ADV, 2 participants in HI, and 9 participants in LI who made constant errors. Again, these numbers suggest large individual differences within proficiency groups. Although the LI group indicated lower mean accuracy than the other two groups, 11 out of 20 participants in the group completed the non-reversible scrambled sentences in an OSV form seven times or more, demonstrating a good understanding of the scrambling structure. On the other hand, there were four participants in the LI group (NNS05, 36, 41, and 55) who never filled in the parentheses in such a way that the completed sentence had a scrambling structure. The first three of them did not form any scrambled sentences in the reversible condition either, which suggests that these participants did not have a grammatical knowledge of scrambling.

The marking of the accusative NP with the "*-ni*" particle, which was quite frequent across the different proficiency groups with the reversible scrambled sentences, was greatly reduced when the participants completed the non-reversible sentences. There were only four participants who used the *ga-ni* combination in completing the non-reversible canonical sentences (cf. Table 2.2). This seems to suggest that the participants were more familiar with the verbs used in the non-reversible condition than those in the reversible condition.

As with the reversible scrambled sentences, the raw frequencies of the cases in which the first NP was marked with a subject marker and the second NP was marked with an object marker were counted for the purpose of examining whether the errors in completing the scrambled sentences were caused by the marking of the first NP with *-ga* or the marking of the second NP with an object marker. The *–ga* marker was counted as the subject marker and *–o* and *–ni* were counted as the object markers. There were 60 instances of wrong NP1*-ga* and 83 instances of wrong NP2*-o*/*ni*. A Chi-square test indicated that the differences in observed frequencies approached significance ($\chi^2(1) = 3.699$ $p = 0.054$), indicating that the participants tended to supply *–o* or *–ni* for the second NP more frequently than *–ga* for the first NP.

2.6.3.3 Locative sentences

Table 2.3 presents the raw frequencies of the particles that each participant used to complete the locative scrambled sentences. There were six opportunities for each participant. The errors that the participants made in this condition appear quite different from the errors in the reversible and non-reversible scrambled sentences. Many of the errors in the reversible and non-reversible scrambled sentences were based on the application of the SOV form where the *–ga* particle was supplied after NP1 and some other particle after NP2. On the other hand, there was only one instance in which –*ga* was supplied after NP1 in the locative scrambled sentences.

Table 2.3: Raw frequencies of the particles used for the locative sentences in Task 1

#	Groups	*de-ga*	*(*)ni-ga*	**de-o*	**ni-o*	#	Groups	*de-ga*	*(*)ni-ga*	**de-o*	**ni-o*	other
NNS02	ADV	5	1			NNS30	HI	4	2			
NNS03	ADV	2	4			NNS31	HI	4	2			
NNS08	ADV	5	1			NNS32	HI	4	2			
NNS10	ADV	6				NNS33	HI	4	2			
NNS11	ADV	6				NNS37	HI		1	3		*de-niX2*
NNS18	ADV	1		5		NNS38	HI	6				
NNS24	ADV	6				NNS43	HI	6				
NNS26	ADV	6				NNS04	LI	1		4		*de-ni*
NNS27	ADV	6				NNS05	LI			5		*de-ni*
NNS29	ADV	6				NNS06	LI	4		2		
NNS40	ADV	6				NNS07	LI	6				
NNS42	ADV	1	4		1	NNS13	LI		6			
NNS44	ADV	5	1			NNS15	LI	6				
NNS45	ADV	6				NNS16	LI	3	3			
NNS46	ADV	5	1			NNS17	LI		3	1	2	
NNS47	ADV	6				NNS21	LI	4		2		
NNS48	ADV	3	3			NNS23	LI	4		2		
NNS52	ADV	3	3			NNS28	LI	5	1			
NNS53	ADV	6				NNS34	LI	6				
NNS54	ADV	6				NNS35	LI	5		1		
NNS01	HI	5		1		NNS36	LI			5		*ga-o*
NNS09	HI	4	2			NNS39	LI	5		1		
NNS12	HI	6				NNS41	LI			1	4	*de-ni*
NNS14	HI	4	2			NNS49	LI			6		
NNS19	HI	6				NNS50	LI	4	2			
NNS22	HI	3	1	1	1	NNS51	LI	1		4	1	
NNS25	HI	5	1			NNS55	LI				1	*ni-deX5*

Note: "NP1-*ni* NP2-*ga*" was considered correct in the "lax" scoring while it was considered incorrect in the "strict" scoring.

The most frequent type of error with the locative scrambled sentences was the "*ni-ga*" marking in which the particle –*ni* was used instead of the correct –*de* particle to mark the locative NP (48 instances). The second most frequent type of error was the "*de-o*" marking in which the participants marked the nominative NP with the *–o* particle (44 instances). While the errors in the choice of *–ni* or *–de* particle to mark locative NPs are quite common among L2 learners of Japanese, the frequent marking of the nominative NP with the *–o* particle is rather mysterious. As there were only 6 participants who marked the nominative NPs with *–o* at least once in the locative canonical sentences but there were 18 participants who supplied *–o* after the nominative NPs in the locative scrambled sentences, the frequent marking of the nominative NPs with *–o* in the locative scrambled sentences could be interpreted as a rather strong manifestation of the L2 learners' tendency to attach *–o* to the preverbal NP.

Figure 2.7 is a histogram showing how many participants in each proficiency group completed the locative scrambled sentences correctly (based on the "lax" scoring).

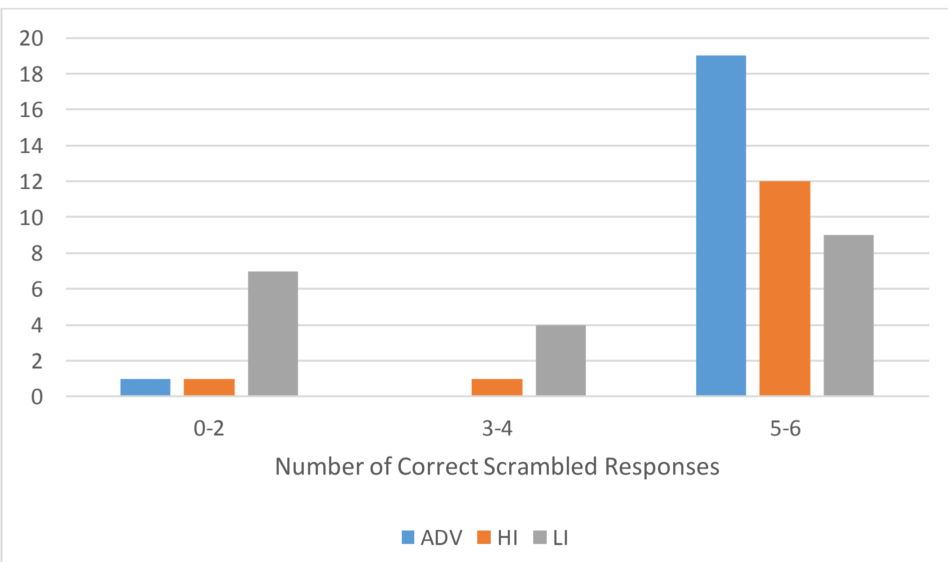


Figure 2.7: The number of participants by proficiency group who completed the locative scrambled sentences correctly in Task 1

If "constant error" is defined as completing the items in a scrambled word order less than three times (i.e., 50%) or fewer, there were 1 participant in ADV, 1 participant in HI, and 8 participants in LI who made constant errors. Although the LI group indicated lower mean accuracy than the other two groups, 12 out of 20 participants in the group completed the non-reversible scrambled sentences in a scrambled word order four times or more, demonstrating a good understanding of the scrambling structure. On the other hand, there were five participants in the LI group (NNS05, 36, 41, 49, and 55) who never filled in the parentheses in such a way that the completed sentence had a scrambling structure. Again, the first three of them did not form any scrambled sentences in the other conditions either, suggesting that these participants did not have a grammatical knowledge of scrambling.

As with the other two sentence types, the raw frequencies were counted where the first NP was marked with a subject marker and the second NP was marked with an object marker. The particle – *ga* was counted as the subject marker. Because there was only one participant who marked NP2 with –*de*, the particles –*o* and –*ni* were counted as the object markers. There was only one instance of wrong NP1-*ga* and there were 60 instances of wrong NP2-*o*/*ni*. A Chi-square test indicated that the difference in observed frequencies was highly significant ($\chi^2(1) = 57.066$, $p < 0.001$).

2.6.4 Discussion

The fill-in-the-blank task was conducted for the purpose of assessing L2 learners' grammatical knowledge of scrambling. As for the relation of L2 learner's general proficiency in Japanese and their grammatical knowledge of scrambling, the analyses of the mean accuracy rate data clearly indicated that there is a positive relation between the two. The "lax" analysis indicated that the LI group was significantly less accurate in completing the scrambled sentences than the other two groups. At the same time, when the raw frequencies of the particles that each participant used were counted, it was suggested that the grammatical knowledge of scrambling and case markers does not necessarily develop uniformly among L2 learners. There were two participants in the ADV group who formed OSV sentences in less than 50% of the opportunities. Also, while the LI group on average was less accurate in completing the OSV sentences, about half of the LI participants constantly demonstrated their grammatical knowledge of scrambling.

The mean accuracy rates of the reversible sentences were constantly lower than those of the non-reversible sentences across proficiency groups. Because the sentence stimuli were strictly separate between the reversible and non-reversible conditions, the present study is unable to separate the effect of reversibility from other causes. It is suspected that the relative unfamiliarity of the reversible verbs also contributed to the lower accuracy rates. Although both types of verbs were taken from the first and second year Japanese textbooks, the reversible verbs chosen (e.g., bully, hit, scold) were intuitively more advanced than the non-reversible verbs (e.g., eat, read, write). In fact, the raw frequency data indicated that the *–ni* particles, instead of the correct *–o* particles, were used quite frequently to mark the accusative NPs in the reversible sentences, suggesting that the participants might not have been as familiar with the reversible verbs. Also, as evident from the comparison of Figures 2.1 and 2.3, the mean accuracy rates of the reversible sentences were constantly lower in the "strict" analysis across the proficiency groups, especially with the HI and LI groups.

The "lax" analysis of the accuracy rates did not observe any significant interaction between Reversibility and Word order. This indicates that, although the items in the reversible condition elicited more errors in general, reversibility on average did not affect the accuracy in completing scrambled sentences. The result was as expected because the grammatical knowledge of scrambling should not be influenced by the semantic contrasts of the NPs in a sentence.

In Task 1, the locative sentences were included for exploratory purposes to examine whether the participants would be more familiar with the scrambling of such sentences than of transitive sentences. It was based on the intuitive assumption that the fronting of the locative NPs in locative sentences are more frequent than the fronting of the accusative NP in transitive sentences. The analyses of the mean accuracy rates indicated that the locative scrambled sentences were no more familiar to the participants than the transitive sentences. The mean accuracy rates of the transitive and locative sentences indicated a similar pattern in which the LI group was significantly less accurate in completing sentences in the scrambled word order.

2.7 Method for Task 2

Task 2 (picture description task) was conducted to assess L2 learners' production performance of scrambled sentences.

2.7.1 Participants

The same 54 NNS participants from Task 1 completed Task 2. The 20 NS participants also completed Task 2 as a comparison group.

2.7.2 Materials and design

As with Task 1, there were four conditions in Task 2: reversible canonical/scrambled and non-reversible canonical/scrambled. Locative sentences were not included in Task 2. Using the same 16 verbs as in Task 1 (eight verbs that can be used reversibly and eight verbs that can be used non-reversibly), 16 each of reversible and non-reversible sentences were made. All the verbs and inanimate nouns were again chosen from *Nakama 1 & 2 2nd Edition* (Hatasa, Hatasa,

& Makino, 2009). Half of the animate nouns were also chosen from the textbooks, but the other half were common Japanese names for the purpose of minimizing the repetition of the same nouns in the item set. As with Task 1, one of the nouns in each reversible sentence was a regular noun (such as "younger brother" and "policeman") while the other noun was a name. In order to alleviate the effect of this regular noun/name contrast on the participants' performance, regular nouns were nominative-marked in half of the reversible sentences and names were nominative-marked in the other half. Half of the non-reversible sentences had regular nouns as their subjects, and the other half had names as subjects.

For each of the reversible and non-reversible sentences, corresponding "reversible" and "non-reversible" pictures were drawn by an artist. The reversible pictures depicted situations in which a person is doing something to another person, and the non-reversible pictures depicted situations in which a person is doing something to an inanimate object. In order to make clearer which of the two persons is the agent of the action in the reversible pictures, the agent was drawn larger and closer to the viewer than the patient. See Appendix 2B for the complete set of the target pictures.

Several pilot trials indicated that it was quite challenging for some L2 learners to describe the pictures, as they were often not sure which vocabulary items to use. There were also quite a few descriptions which deviated from the sentence structure that the present task wished to elicit (i.e., simple transitive sentences). In order to alleviate these problems, the persons and the inanimate objects in the pictures were labeled with the words in the original sentences. As it was expected that the arrangement of the words might influence the participants' productions of the pictures, the labels for the agent and patient were placed at the same height in each picture. Also, the label for the agent was placed on the left side for half of the pictures and on the right side for the other half. The label for the verb was placed at the bottom of each picture.

In order to elicit canonical and scrambled sentences, canonical and scrambled versions were made for each picture. In the canonical version, the label for the agent was highlighted in orange, and in the scrambled version, the label for the patient was highlighted. Figure 2.8 shows examples of the reversible canonical/scrambled pictures and the non-reversible canonical/scrambled pictures.

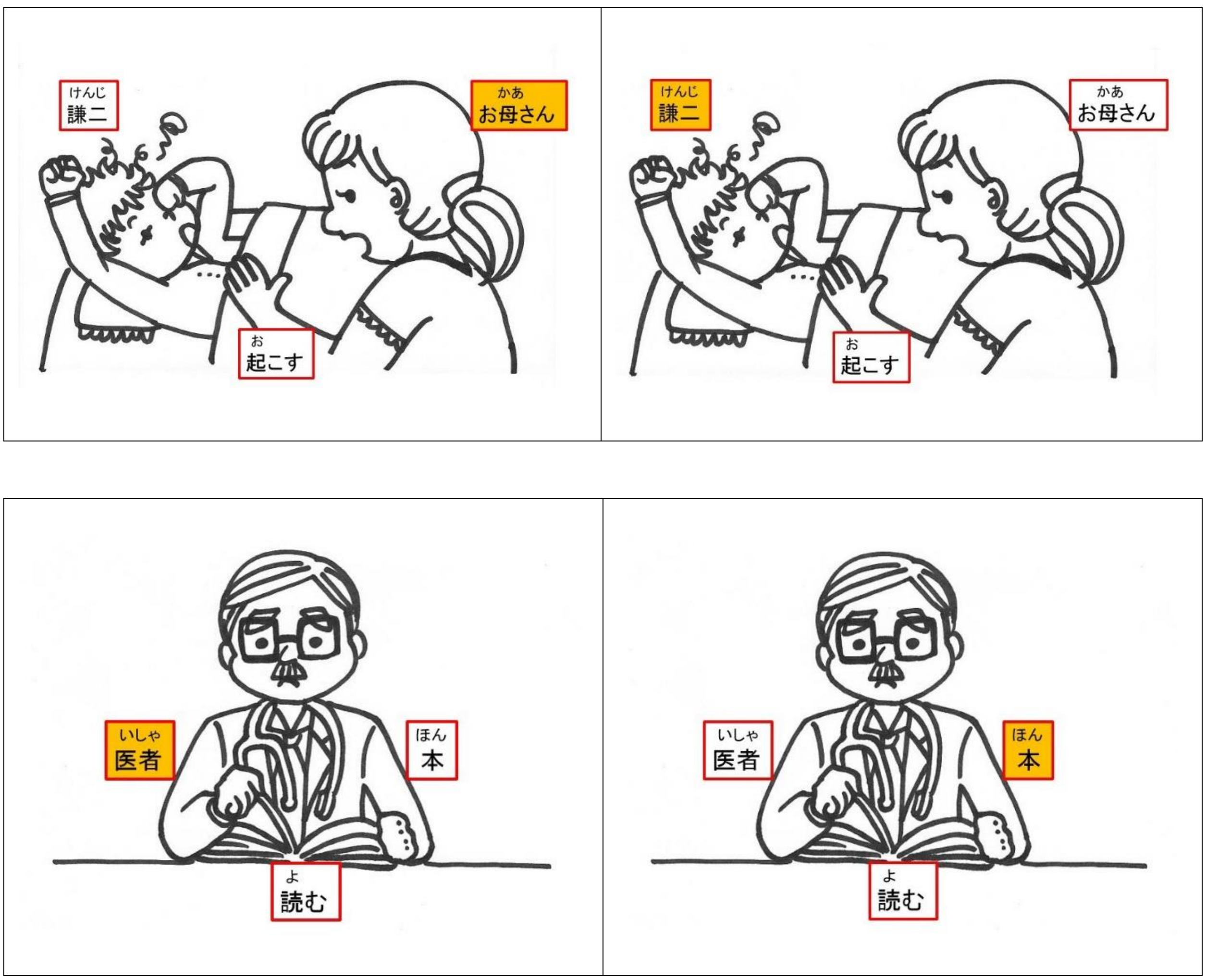


Figure 2.8: Examples of the pictures used in Task 2: reversible canonical/scrambled (above) and non-reversible canonical/scrambled (below)

The canonical/scrambled pairs of the pictures were distributed into two item lists so that a participant would not see the same picture during the task. Each item list consisted of 8 each of reversible canonical/scrambled pictures and non-reversible canonical/scrambled pictures as well as 48 filler pictures, totaling 80 pictures. The filler pictures were based on different sentence types (e.g., locative, ditransitive), and the label for the agent was highlighted in all of the filler pictures. The labels for the agent were placed on the left for half of the filler pictures, and they were placed on the right for the other half. The pictures in each list were pseudo-randomized so that the pictures to elicit the same sentence type would not be adjacent to each other. In an effort to alleviate the influence of the item presentation order, two pseudo-randomized versions were created for each item list.

2.7.3 Procedures

Prior to the task, the participants were given a word list and were instructed to study the words in the list. The word list included all of the vocabulary items that were in the target pictures and about half of the vocabulary items that were in the filler pictures. Glosses in English were provided next to each Japanese word. In addition to checking the meaning of the words, the NNS participants were also asked to pronounce each word aloud, while the NS participants were only asked to study the words in the list. The purpose of the word list was to alleviate the effect of unfamiliar words or unfamiliar pronunciation of the words on the participants' performance during picture description.

The presentation of the pictures was carried out on a laptop computer with Microsoft PowerPoint. Each trial began with the participant pressing the right arrow key, which displayed a picture with word labels (without case particles) on the computer monitor and played a beep sound simultaneously. After one second, the software automatically displayed the next slide in which one of the word labels was highlighted in orange. The participant described the picture starting with the highlighted word, and moved onto the next item by pressing the right arrow key. There were designated short rest periods after each 20 trials. The participants' oral responses were recorded with digital recorders.

The participants were instructed to describe the picture in a full sentence, starting with the highlighted word and using all the words in the picture. They were also instructed to start the description as soon as possible after the beep sound. However, they were asked not to say the highlighted word first and then think about how to formulate the sentence. They were asked, instead, to formulate the sentence first and then say the sentence aloud as soon as possible. Eight practice items preceded the study items. The use of the topic marker –*wa* was not discouraged in the present task for the purpose of encouraging natural sentence production.

2.7.4 Scoring and analysis

2.7.4.1 Mean response accuracy data

The participants' oral responses to each item were scored for their accuracy. Responses were given the score of 1 if they were correct, and 0 if they were incorrect. As with Task 1, responses were judged correct only if both of the particles in a sentence were correct. If the participant self-corrected after one or more incorrect particles were used, the response was given a score of 0.5. On the other hand, if the participant initially used one or more correct particles and then used wrong particles to revise the response, a score of 0 was given. Unintended responses such as passives, relative clauses, and failure to start the description with the highlighted word (all of which took place only with the scrambled pictures) were treated as missing data.

The topic marker –*wa* was treated as equivalent to the subject marker –*ga* in scoring, as –*wa* often signifies the agent of the action (although it is a topic marker) especially in simple transitive sentences such as those used in the present study. Therefore, the correct combinations of particles were *ga*/*wa*-*o* (i.e., NP1-*ga* or –*wa* AND NP2-*o*) for the canonical sentences and *o*-*ga*/*wa* for the scrambled sentences. The omission of the accusative case marker in canonical sentences (i.e., *wa*/*ga*-*Ø*) was considered correct. In addition, the *wa*-*ga* combination was also treated as a correct response for the scrambled sentences because the topic marker –*wa* can also topicalize the accusative NP in a sentence.

As with Task 1, the NNS participants frequently used the –*ni* particle to mark accusative NPs instead of the correct –*o* particle, especially for the reversible sentences. Therefore, another set of scoring was conducted which included the –*ni* marking of the direct object as a correct answer. In this second scoring, the use of a particle other than –*o* or –*ni* to mark the accusative NP was also considered correct if the same particle was used to mark the accusative NP of the canonical and scrambled sentences that shared the same verb. Otherwise, such responses were given the score of 0. Again, the first scoring will be called the “strict” scoring and the second scoring the “lax” scoring for ease of explanation. The results based on the “lax” scoring will be presented first as this scoring appears to better reflect the participants’ performance on the production of scrambled sentences.

The mean accuracy data were analyzed using linear mixed-effects modeling in R. The accuracy data for the reversible and non-reversible sentences were analyzed first by proficiency group, using Reversibility and Word order as fixed effects. For the purpose of analyzing the differences among the proficiency groups, a subsequent analysis was conducted with the addition of Proficiency as a fixed effect. Subjects and Items were random effects. As the NS participants did not participate in Task 1 (the fill-in-the-blank task) and their responses in Task 2 were somewhat different. For this reason, the data from the NS participants will be analyzed separately from those of the NNS participants.

2.7.4.2 Raw frequency data

As with Task 1, raw frequencies of the particles that each participant used to produce the scrambled sentences were counted for the purpose of observing response patterns of individual participants. The frequencies of "NP1-*ga/wa*" and "NP2-*o/ni*" in the items of the scrambling condition were also counted for the purpose of observing whether there was a greater tendency to add *–ga or –wa* erroneously after the first NPs or to add *–o* or *–ni* after the second NPs.

2.7.4.3 Reaction time data

The reaction time (RT) data were collected by measuring the time between the onset of the beep sound and the onset of the picture description. If the participant self-corrected the response, the time between the beep sound and the onset of the self-correction was recorded as the RT for the response. The sound files of each participant's responses were imported into sound analyzing software (a non-commercial edition of *WavePad Sound Editor* by NCH Software), and the RTs were manually measured. Only correct responses based on the "lax" scoring were included in the RT analysis. Note, however, that although the software allowed measurements to a millisecond level and the researcher made his best efforts to capture the RT as accurately as possible, the RT data were not accurate to a millisecond level in the end. This is because the onset of the sentence production was sometimes not clear especially when the initial word started with a voiceless fricative sound. Therefore, the RT data should be interpreted with caution.

A visual inspection of the raw RT data indicated a positive skew. Therefore, the RTs were transformed with a reciprocal transformation before they were analyzed using the linear mixed-effects modeling. The transformed RT data were analyzed by proficiency group, using Reversibility and Word order as fixed effects. Subjects and Items were random effects.

2.8 Results of Task 2

2.8.1 Mean response accuracy data based on the "lax" scoring

2.8.1.1 NNS participants

Figure 2.9 shows the mean accuracy rates of the reversible and non-reversible sentences by proficiency level based on the "lax" scoring.

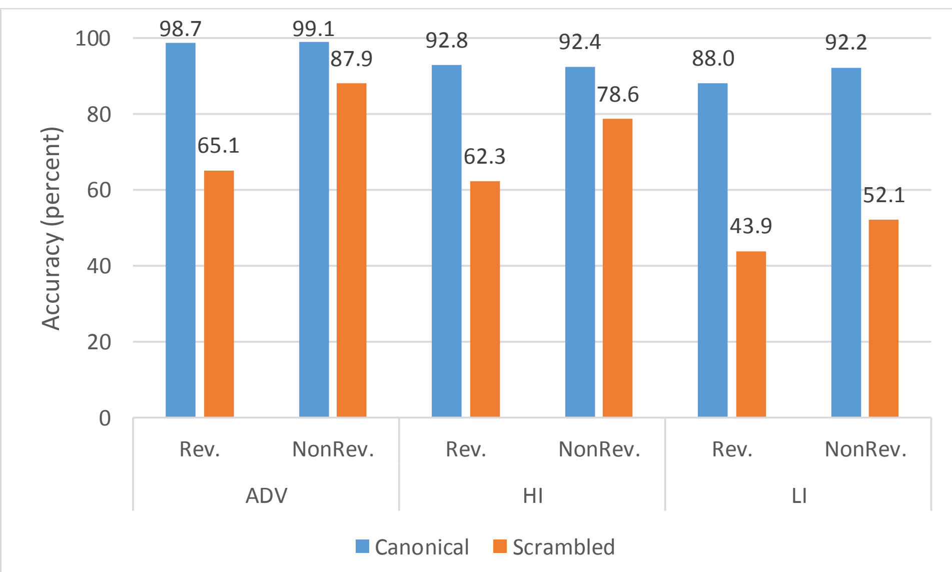


Figure 2.9: Mean accuracy rates by the NNS groups based on the "lax" scoring in Task 2

The data of the ADV and HI groups indicated a significant interaction of Reversibility and Word order. Therefore the following model was used for analyses: `Accuracy ~ Reversibility * Word order + (1 | subject) + (1 | item)`. On the other hand, the data of the

LI group did not indicate a significant interaction of the fixed effects, and therefore the following model was used to analyze the data: `Accuracy ~ Reversibility + Word order + (1 | subject) + (1 | item).`

For the ADV group, the main effects of both Reversibility and Word order were significant (Reversibility: $t = 4.967$, $p < 0.0001$; Word order: $t = 10.98$, $p < 0.0001$). The interaction of the two factors was also significant ($t = 5.529$, $p < 0.0001$). The simple effect of Word order was significant for the reversible sentences ($t = 11.764$, $p < 0.0001$) as well as for the non-reversible sentences ($t = 4.35$, $p < 0.0001$). On the other hand, an analysis using a model with random slopes (by-subject and by-item for the effect of Word order) indicated a substantial decrease in the effect of Word order for both the reversible and non-reversible sentences (Reversible: $t = 6.189$, $p < 0.0001$; Non-reversible: $t = 2.413$, $p = 0.0227$). All in all, while the ADV participants were less accurate in producing scrambled sentences both for reversible and non-reversible sentences, they seem to have experienced more difficulties with the reversible scrambled sentences.

For the HI group, the main effect of Word order was significant ($t = 6.699$, $p < 0.0001$) while the main effect of Reversibility only approached significance ($t = 1.987$, $p = 0.0561$). The interaction of the two factors was significant ($t = 2.524$, $p = 0.012$). The simple effect of Word order was significant for the reversible sentences ($t = 6.542$, $p < 0.0001$) as well as for the non-reversible sentences ($t = 3.033$, $p = 0.0026$). As with the ADV group, an analysis using a random slopes model (by-subject for the effect of Word order) indicated a decrease in the effect of Word order for both the reversible ($t = 4.808$, $p < 0.0001$) and non-reversible sentences ($t = 2.251$, $p = 0.0348$). Taken together, the result indicated that the HI participants, similarly to the ADV participants, had more difficulty with the reversible scrambled sentences than the non-reversible counterparts although the interaction was not as robust as that of the ADV group.

For the data of the LI group, the main effect of Word order was significant (t = 14.944, $p <$ 0.0001), and the main effect of Reversibility was also significant (t = 2.213, p = 0.0273). The interaction of the two factors was not significant ($\chi^2(1) = 0.5436$, p = 0.461). An analysis using a random slopes model (by-subject for the effect of Word order) indicated a substantial decrease in the effect of Word order (t = 4.64, p = 0.0002). The results thus indicated that the LI participants were equally less accurate with the reversible and non-reversible scrambled sentences as compared to their canonical counterparts.

In order to compare the differences among the three proficiency groups, an analysis including Proficiency as the third fixed effect was carried out. A model with random slopes was used in this analysis. The following linear mixed-effects model was justified by the data: `Accuracy ~ Reversibility * Word order * Proficiency + (1 + Word order | subject) + (1 + Word order | item)`. The main effect of Word order was highly significant (t = 6.764, $p <$ 0.0001). The main effect of Reversibility was also significant (t = 2.339, p = 0.0222). The effect of Proficiency was significant between the ADV and LI groups (t = 3.033, p = 0.0038), but it was not significant either between the ADV and HI groups (t = 1.444, p = 0.1548) or between the HI and the LI groups (t = 1.308, p = 0.1966). The interaction of Reversibility and Word order was significant ($\chi^2(1) = 12.099$, p = 0.0005), but neither the interaction of Reversibility and Proficiency ($\chi^2(2) = 1.7299$, p = 0.4211) nor the interaction of Word order and Proficiency ($\chi^2(2) = 4.4267$, p = 0.1093) was significant. The three-way interaction among the fixed effects was significant ($\chi^2(2) = 8.2253$, p = 0.0164).

The pairwise comparison of the sentence types between the proficiency groups indicated that, for the reversible canonical sentences, the only difference in accuracy between the ADV and LI groups was significant (t = 2.348, p = 0.0212). For the reversible scrambled sentences, the differences between the ADV and LI groups as well as between the HI and LI groups approached significance (ADV-LI: t = 1.951, p = 0.0559; HI-LI: t = 1.761, p = 0.0836), but the difference between the ADV and HI groups was not significant. For the non-reversible canonical sentences, none of the differences between the groups reached a level of significance. For the non-reversible scrambled sentences, the differences between the ADV and LI groups as well as between the HI and LI groups were significant (ADV-LI: t = 3.488, p = 0.0009; HI-LI: t = 2.498, p = 0.0154), but the difference between the ADV and HI groups was not significant.

The pairwise comparison of the interactions between the proficiency groups indicated that, for the reversible sentences, the interaction of Word order and Proficiency was not significant between either of the proficiency groups. On the other hand, for the non-reversible sentences, the interaction of Word order and Proficiency was significant between the ADV and the LI groups (t = 2.516, p = 0.0145) and between the HI and the LI groups (t = 2.265, p = 0.027), but not between the ADV and the HI groups (t = 0.021, p = 0.9831). The interaction of Reversibility and Proficiency for the scrambled sentences was significant between the ADV and LI groups (t = 2.992, p = 0.0028). The three-way interaction was significant only between the ADV and the LI groups (t = 2.803, p = 0.0051). No other interactions were significant. One difference between this three-factor analysis and the previous by-group analyses is that the simple effect of Word order for the non-reversible sentences in the ADV and HI groups did not reach a level of significance in this three-factor analysis (ADV: t = 1.959, p = 0.0542; HI: t = 1.685, p = 0.0968).

Taken together, the results indicated that all three of the NNS groups were less accurate in producing the reversible scrambled sentences as compared to the canonical counterparts. On the other hand, for the non-reversible sentences, while all three of the groups were also less accurate in producing the OSV word order, the LI group was significantly less accurate in their production of the scrambled sentences.

2.8.1.2 NS participants

Figure 2.10 shows the NS group's mean accuracy rates for the reversible and non-reversible sentences based on the "lax" scoring. The following model was used to analyze the data:

```
Accuracy ~ Reversibility * Word order + (1 | subject) + (1 |
item).
```

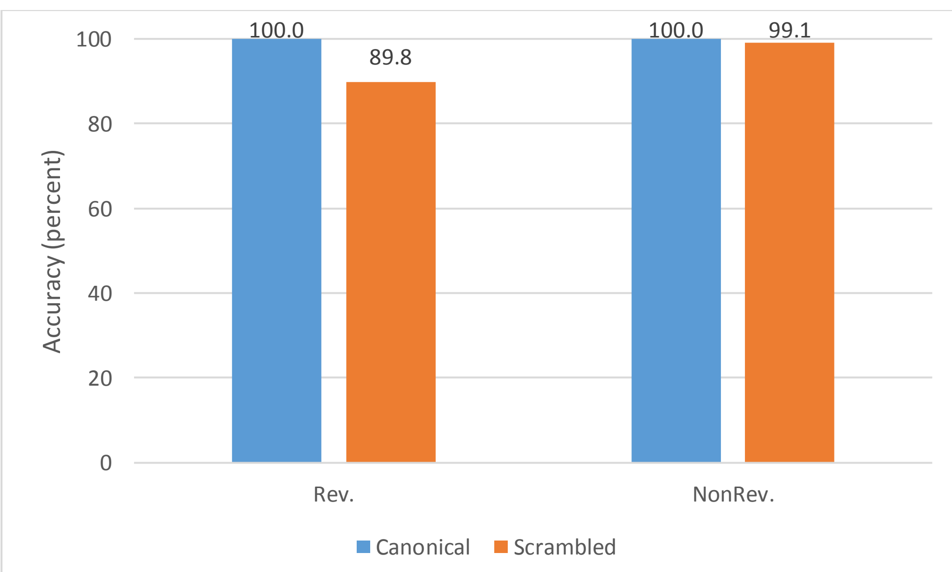


Figure 2.10: Mean accuracy rates by the NS group based on the "lax" scoring in Task 2

As the model suggests, the interaction of Reversibility and Word order was significant (t = 4.404, p < 0.0001). The simple effect of Word order was highly significant for the reversible sentences

($t = 6.188$, $p < 0.0001$), but it was not significant for the non-reversible sentences ($t = 0.681$, $p = 0.496$). However, when the same dataset was analyzed using a model with random slopes (by-subject and by-item for the effect of Word order), the interaction of the two factors was not significant ($t = 1.756$, $p = 0.0907$), while the simple effect of Word order for the reversible sentences remained significant ($t = 2.449$, $p = 0.0194$). The results thus indicated that the NS participants' production of reversible scrambled sentences was less accurate than the canonical counterparts. However, this seemingly lower accuracy with the reversible scrambled sentences should be interpreted with caution. There were only four instances (out of 160 opportunities) of the errors based on the SOV form when the NS participants were prompted to produce reversible scrambled sentences. The error rate appears higher because, as will be discussed later, there were a large number of unintended forms when the NS participants described the scrambled pictures, and such unintended sentences were treated as missing data points.

2.8.2 Mean response accuracy data based on the "strict" scoring

2.8.2.1 NNS participants

This section will present the analyses of the accuracy data based on the "strict" scoring, in which the only correct combinations of particles were *ga/wa-o/Ø* for the canonical sentences, and *o-ga/wa* and *wa-ga* for the scrambled sentences. Different mixed-effects models were used to analyze the data of the three proficiency groups. Figure 2.11 shows the mean accuracy rates by proficiency group.

The data of the ADV and HI groups did not indicate a significant interaction of Reversibility and Word order. Therefore the following model was used for analyses: `StrictAccuracy ~ Reversibility + Word order + (1 | subject) + (1 | item)`. On the other hand, the data of the LI group indicated a significant interaction of the fixed effects, and therefore the following model was used to analyze the data: `Accuracy ~ Reversibility * Word order + (1 | subject) + (1 | item)`.

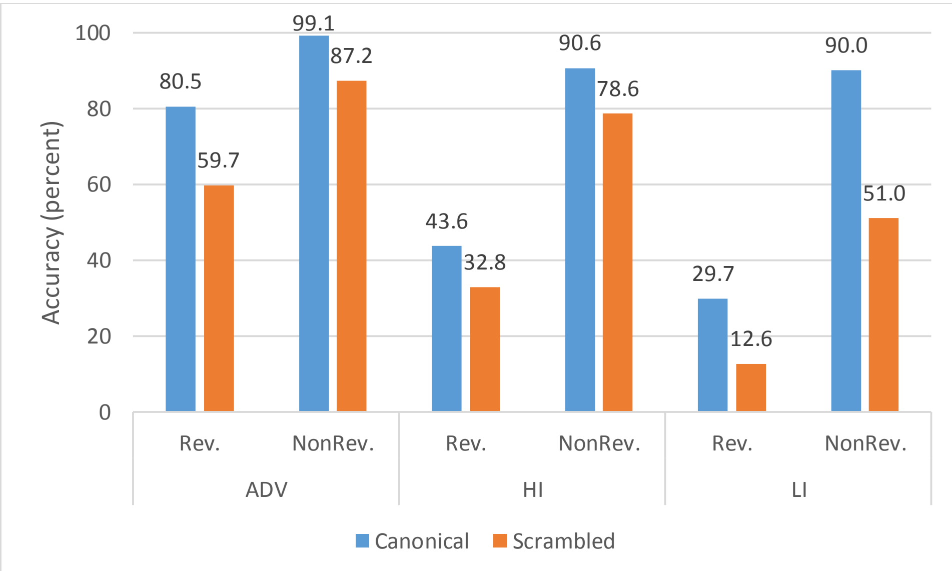


Figure 2.11: Mean accuracy rates by the NNS groups based on the "strict" scoring in Task 2

For the ADV group, the main effects of Reversibility as well as Word order were significant (Reversibility: $t = 8.99$, $p < 0.0001$; Word order: $t = 7.197$, $p < 0.0001$). The interaction of the two factors did not reach a level of significance ($\chi^2(1) = 2.6125$, $p = 0.106$). An analysis using a random slopes model (by subject for the effect of Reversibility and Word order) indicated a decrease in the two fixed effects while both remained highly significant (Reversibility: $t = 5.423$, $p < 0.0001$; Word order: $t = 3.937$, $p = 0.0009$).

For the HI group, the main effects of Reversibility as well as Word order were significant (Reversibility: $t = 7.96$, $p < 0.0001$; Word order: $t = 2.96$, $p = 0.0033$). The interaction was not significant ($\chi^2(1) = 0.0683$, $p = 0.7938$). An analysis with a random slopes model (by-subject for the effect of Reversibility and Word order) indicated a decrease in the two fixed effects. While the main effect of Reversibility remained highly significant ($t = 4.884$, $p < 0.0001$), the main effect of Word order only approached significance in this analysis ($t = 1.808$, $p = 0.0972$).

For the LI group, as the model suggests, the interaction of Reversibility and Word order was significant ($t = 3.711$, $p = 0.0002$), where the difference in the accuracy rates between the canonical and scrambled sentences was larger for the non-reversible sentences than the reversible sentences. For the reversible sentences, the simple effect of Word order was significant ($t = 4.077$, $p < 0.0001$), and the simple effect of Word order was highly significant for the non-reversible sentences ($t = 9.235$, $p < 0.0001$). An analysis with a random slopes model (by-subject for the effects of Reversibility and Word order) indicated a decrease in the effect of Word order for both the reversible sentences ($t = 2.355$, $p = 0.0269$) and the non-reversible sentences ($t = 5.245$, $p < 0.0001$).

As with the "lax" data, an analysis including Proficiency as the third fixed effect was carried out for the purpose of observing the differences among the three NNS groups. Due to the variability of subjects observed in the analyses above, a model with random slopes was used in this analysis. The following linear mixed-effects model was justified by the data: `StrictAccuracy ~ Reversibility * Word order * Proficiency + (1 + Reversibility + Word order | subject) + (1 | item)`. The main effects of Reversibility and Word order were significant (Reversibility: $t = 7.649$, $p < 0.0001$; Word order: $t = 5.33$, $p < 0.0001$). The effect of Proficiency was significant between the ADV and HI groups ($t = 4.138$, $p = 0.0001$) as well as between the ADV and LI groups ($t = 6.43$, $p < 0.0001$), but it did not reach a level of significance between the HI and LI groups ($t = 1.697$, $p = 0.0958$). The interaction of Reversibility and Proficiency was significant ($\chi^2(2) = 13.369$, $p = 0.0016$), but the interaction of Word order and Proficiency only approached significance ($\chi^2(2) = 5.6933$, $p =$

0.058). The interaction of Reversibility and Word order was not significant ($\chi^2(1) = 1.6086$, $p = 0.2047$). The three-way interaction of the fixed effects was significant ($\chi^2(2) = 19.494$, $p < 0.0001$).

The pairwise comparisons of each sentence type across proficiency groups indicated that, for the reversible canonical sentences, there were significant differences in accuracy between the ADV and HI groups as well as between the ADV and LI groups (ADV-HI: $t = 4.261$, $p < 0.0001$; ADV-LI: $t = 6.515$, $p < 0.0001$), but the difference between the HI and LI groups was not significant ($t = 1.649$, $p = 0.1048$). For the reversible scrambled sentences, the difference in accuracy was significant between the three groups (ADV-HI: $t = 2.258$, $p = 0.0279$; ADV-LI: $t = 4.793$, $p < 0.0001$; HI-LI: $t = 2.109$, $p = 0.0396$). For the non-reversible canonical sentences, none of the differences in accuracy was significant among the three proficiency groups although the difference between the ADV and LI groups approached significance ($t = 1.833$, $p = 0.0714$). For the non-reversible scrambled sentences, the difference between the ADV and LI groups as well as between the HI and LI groups was significant (ADV-LI: $t = 3.559$, $p = 0.0008$; HI-LI: $t = 2.463$, $p = 0.017$), but the difference between the ADV and HI groups was not significant ($t = 0.775$ $p = 0.4418$).

The pairwise comparisons of the interactions across proficiency groups indicated that, for the canonical sentences, the interaction of Reversibility and Proficiency was significant between the ADV and HI groups as well as between the ADV and LI groups (ADV-HI: $t = 2.837$, $p = 0.006$; ADV-LI: $t = 4.634$, $p < 0.0001$), but it was not significant between the HI and LI groups ($t = 1.367$, $p = 0.1762$). For the scrambled sentences, the interaction of Reversibility and Proficiency was not significant between any of the groups.

For the reversible sentences, none of the interaction of Word order and Proficiency was significant between groups. For the non-reversible sentences, the interaction of Word order and Proficiency was significant between the ADV and the LI groups ($t = 2.733$, $p = 0.008$) and also between the HI and the LI groups ($t = 2.527$, $p = 0.0138$), but it was not significant between the ADV and the HI groups.

Finally, the three-way interaction was significant between the ADV and LI groups ($t = 4.404$, $p < 0.0001$) as well as between the HI and LI groups ($t = 2.367$, $p = 0.018$), but it was not significant between the ADV and HI groups.

The results above indicated that, for the non-reversible sentences, while the LI group was significantly less accurate in producing the scrambled sentences, the accuracy of the canonical sentences was comparable across the proficiency groups. For the reversible sentences, on the other hand, the HI and LI groups were significantly less accurate than the ADV group with the canonical sentences, and the three groups performed differently from each other for the scrambled sentences. The effect of Word order for the reversible sentences was greater for the ADV group than for the HI group, but it was possibly due to the higher accuracy rate of the ADV group with the reversible canonical sentences.

2.8.2.2 NS participants

Figure 2.12 presents the NS group's mean accuracy rates of the reversible and non-reversible sentences based on the "strict" scoring. The following model was used to analyze the data: `StrictAccuracy ~ Reversibility * Word order + (1 | subject) + (1 | item)`.

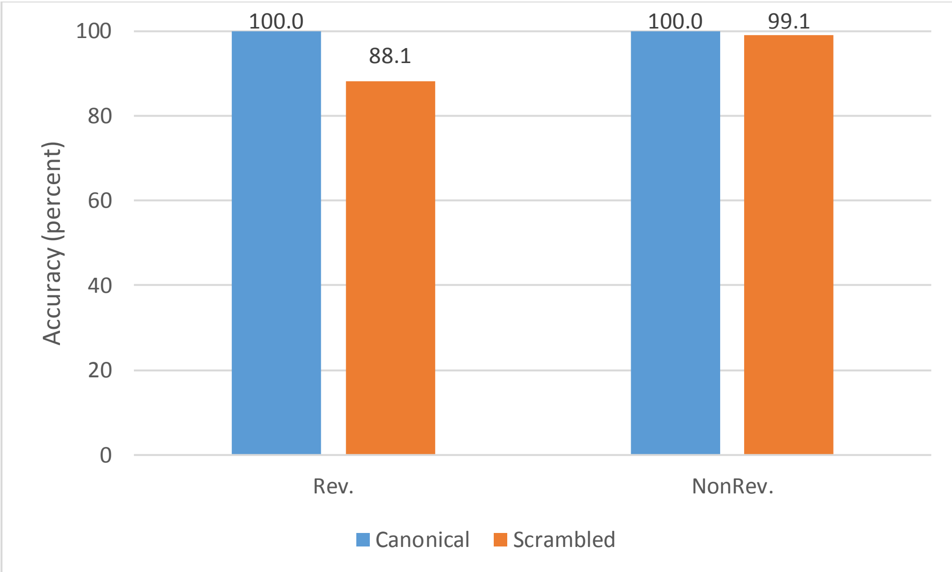


Figure 2.12: Mean accuracy rates by the NS group based on the "strict" scoring in Task 2

As the model suggests, there was a significant interaction of Reversibility and Word order (t = 4.904, $p < 0.0001$). The simple effect of Word order was significant for the reversible sentences (t = 6.820, $p < 0.0001$), but it was not significant for the non-reversible sentences (t = 0.641, p = 0.522). The result again needs to be interpreted with caution, however, due to the rather large amount of missing data (i.e., productions of non-target sentence types) when the NS participants were prompted to produce reversible scrambled sentences.

2.8.3 Raw frequency data

This section will present analyses based on the raw frequencies of the particles that participants used to produce the reversible and non-reversible scrambled sentences. The purpose of the analyses is to observe response patterns at the individual level. The data from the NNS participants will be studied first followed by the data from the NS participants.

2.8.3.1 NNS participants: Reversible sentences

Table 2.4 presents the raw frequencies of the particles that each participant used to produce the reversible scrambled sentences. The numbers are based on the final self-corrected response, and do not include the responses before self-correction. There were eight opportunities for each participant.

Table 2.4: Raw frequencies of the particles that the NNS participants used to produce reversible scrambled sentences in Task 2

#	Groups	*o-ga/wa*	*(*)ni-ga/wa*	**ga/wa-o*	**ga/wa-ni*	**ga-wa*	**ni-o*	**ga/wa-kara*	Passive	other
NNS02	ADV	8								
NNS03	ADV	4			2			1		*o-ni*
NNS08	ADV	2			1				5	
NNS10	ADV	4			1		1			*o-oX2*
NNS11	ADV	6	1	1						
NNS18	ADV		1	1	6					
NNS24	ADV	6	2							
NNS26	ADV	8								
NNS27	ADV	2			1				4, (1)	
NNS29	ADV	5							2, (1)	
NNS40	ADV	6			1					*o-o*
NNS42	ADV	4			1			1	1	*switch*
NNS44	ADV		2	1					(5)	
NNS45	ADV	8								
NNS46	ADV	1		1	1					*kureruX4, CP*
NNS47	ADV		1		2			1	4	
NNS48	ADV	6	1						1	
NNS52	ADV	3		1	1				3	
NNS53	ADV				1				4	*o-ni, CaX2*
NNS54	ADV	7								*switch*
NNS01	HI	2	2		1	1				*o-Ø, Ø- Ø*
NNS09	HI	3				1	1	1		*o-o, to-o*
NNS12	HI	2	3	1	1			1		
NNS14	HI		1		1				2	*no-ga, RCX3*
NNS19	HI	1	3		1			3		
NNS22	HI		5		1		1			*to-wa*
NNS25	HI	5	2				1			
NNS30	HI	6	2							
NNS31	HI	1		1	4					*o-ni, add*
NNS32	HI	4	2		2					
NNS33	HI	3	1					1		*wa-wa, o-de, o-kara*
NNS37	HI	5			1					*switchX2*
NNS38	HI	1	6							*switch*
NNS43	HI	1	4		1			1		*ni-kara*

Table 2.4: Raw frequencies of the particles that the NNS participants used to produce reversible scrambled sentences in Task 2 (continued)

#	Groups	*o-ga/wa*	*wa-ga*	*(*)ni-ga/wa*	**ga/wa-o*	**ga/wa-ni*	**ga-wa*	**ni-o*	**ga/wa-kara*	Passive	other
NNS04	LI				2	3		1			*o-Ø, to-Ø*
NNS05	LI		1		2	5					
NNS06	LI			7		1					
NNS07	LI	6		2							
NNS13	LI			4				2			*ga-ga, ni-ni*
NNS15	LI	1		7							
NNS16	LI			6		1				(1)	
NNS17	LI				3	3				1	*CP*
NNS21	LI	2	1			1			1	1	*to-o, Ca*
NNS23	LI			3	1	4					
NNS28	LI	2		5		1					
NNS34	LI		1	6			1				
NNS35	LI	2				1			4		*kara-ga*
NNS36	LI		1		3	1	2				*to-o*
NNS39	LI			7			1				
NNS41	LI				1	7					
NNS49	LI				7						*ni-to*
NNS50	LI	2		3		1					*ga-ga, ni-Ø*
NNS51	LI					2		6			
NNS55	LI			5	1			2			

Notes: The numbers in parentheses on the "Passive" column indicates the frequencies of incorrect use of passives. The abbreviations on the "other" column stand for the following: *add*: additional words or expressions during the sentence production; *Ca*: causative; *CP*: causative-passive; *kureru*: -tekureru construction; *RC*: relative clause; *switch*: starting the production with the un-highlighted NP (the agent NP).

Figure 2.13 is a histogram showing how many participants in each proficiency group produced the reversible scrambled sentences correctly (based on the "lax" scoring).[19] As with Task 1, large individual differences can be observed within each participant group.

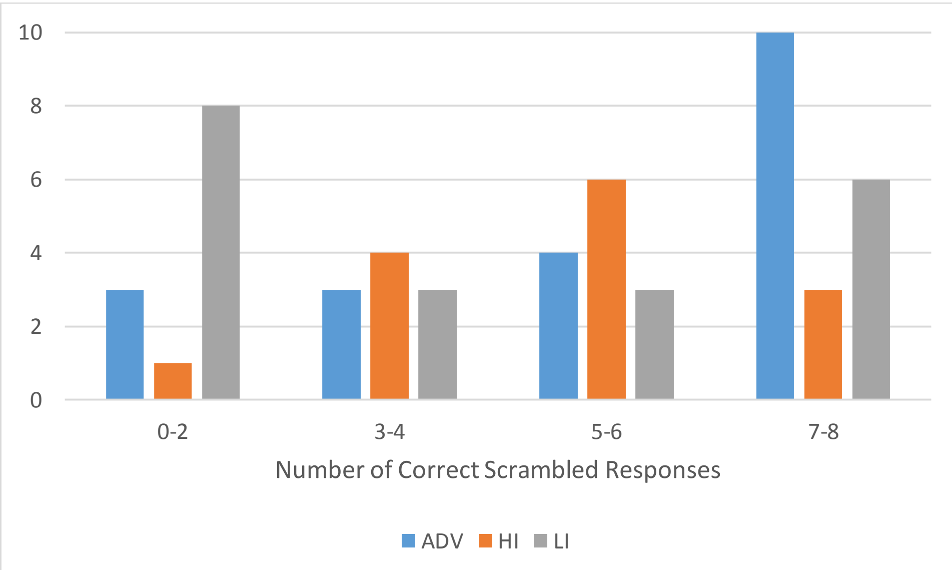


Figure 2.13: The number of participants by proficiency group who produced the reversible scrambled sentences correctly in Task 2

There were 13 ADV participants, 11 HI participants, and 15 LI participants who marked the first NP with *–ga* or *–wa* and the second NP with another particle at least once, suggesting the use of the SOV template when the participants described the reversible scrambled pictures. One participant in the ADV group (NNS53) and five participants in the LI group (NNS04, 17, 41, 49, and 51) never described the pictures using a scrambling structure. While the ADV participant described the pictures correctly using the passive structure four times and also one of the LI participants (NNS17) used passive once, the other four LI participants never described the reversible scrambled pictures correctly. Thirteen participants used (or attempted to use) the passive structure at least once to describe the pictures. Four participants described the pictures

[19] In Figure 2.13, correct responses in passive forms were counted as correct scrambled sentences.

starting with the non-highlighted noun (and therefore such descriptions were in a regular canonical word order). Since such switching was never observed with the canonical pictures, this might also suggest the use of the SOV template during the picture description.

One type of error which was somewhat frequent with the reversible canonical pictures was the marking of the second NP with *–kara* (“from”). The particle *–kara* may be used to mark the agent in the passive construction. Because the second NPs in the scrambled pictures were the agent of the action specified by the verb, it is possible that those participants attempted to deal with the reversible scrambled pictures by marking the agent noun with *–kara*. Alternatively, some of them might have simply forgotten to conjugate the verb into the passive form.

In light of the errors with *–kara*, the high frequency of the “*ga*/*wa*-*ni*” errors in this condition may be explained. There were 17 participants who completed the reversible scrambled sentences with the “*ga*-*ni*” particles in Task 1 whereas there were 33 participants who described the reversible scrambled pictures with the “*ga*/*wa*-*ni*” particles in Task 2. Because the *–ni* particle also may be used to mark the agent noun in the passive sentences, some of the “*ga*/*wa*-*ni*” errors might be partly due to the participants’ conscious efforts to mark the agent with *–ni* in addition to the overuse of the SOV template and/or the simple error in the choice of particles.

For the purpose of examining whether the errors in completing the scrambled sentences were caused by the overuse of subject markers after the first NPs or object markers after the second NPs, the raw frequencies of the cases in which the first NP was marked with a subject marker and the second NP was marked with an object marker were counted. The subject marker was defined as *–ga* or *–wa* and the object markers were *–o* or *–ni* in this analysis. The raw frequencies in this analysis takes self-correction into consideration. That is, if, for instance, a participant first marked the first NP with *–ga* and then corrected it to *–o* in describing a scrambled picture, the score of 0.5 was given to the response for the wrong NP1-*ga* marking. Total scores were 125 wrong NP1-*ga* and 119 wrong NP2-*o*/*ni*. A Chi-square test indicated that there was no statistical differences between the two frequencies ($\chi^2(1) = 0.148$, $p = 0.701$).

2.8.3.2 NNS participants: Non-reversible sentences

Table 2.5 presents the raw frequencies of the particles that each NNS participant used to describe the non-reversible scrambled pictures. There were eight opportunities for each participant. Figure 2.14 is a histogram showing how many participants in each proficiency group produced the reversible scrambled sentences correctly (based on the “lax” scoring).[20]

There were 3 participants in the ADV group, 3 participants in the HI group, and 9 participants in the LI group (a total of 15 participants) who marked the first NP with –*ga* and the second NP with another particle at least once. As will be discussed later, there were also 15 participants in the previous fill-in-the-blank task who marked the first NP with –*ga* and the second NP with another particle at least once in completing non-reversible scrambled sentences. It is thus speculated that the errors in producing non-reversible scrambled sentences were not necessarily caused by the NNV strategy but instead by the lack of grammatical knowledge.

One ADV participant (NNS18), one HI participant (NNS22), and seven LI participants (NNS04, 05, 36, 41, 49, 51, and 55) never produced a scrambled sentence in the description of the non-reversible scrambled pictures. While NNS18 described the pictures correctly four times using the relative clause, the other eight participants never produced a correct description of the non-reversible scrambled pictures.

[20] Correct responses in the form of relative clauses and passives (although somewhat awkward) were counted as scrambled sentences in Figure 2.14.

Table 2.5: Raw frequencies of the particles that the NNS participants used to produce non-reversible scrambled sentences in Task 2

#	Groups	*o-ga/wa*	*wa-ga*	*(*)ni-ga*	**ga/wa-o*	**ga/wa-ni*	**ni-o*	**o-ni*	Passive	RC	Switch	other
NNS02	ADV	8										
NNS03	ADV	7									1	
NNS08	ADV	8										
NNS10	ADV	3								2	3	
NNS11	ADV	8										
NNS18	ADV				2					4, (2)		
NNS24	ADV	3	3							2		
NNS26	ADV	8										
NNS27	ADV	7										*Ø-o*
NNS29	ADV	8										
NNS40	ADV	8										
NNS42	ADV	8										
NNS44	ADV	1		1	3				(2)			*ga-ga*
NNS45	ADV	8										
NNS46	ADV	6	1									*o-o*
NNS47	ADV	7									1	
NNS48	ADV	6										*o-oX2*
NNS52	ADV	8										
NNS53	ADV	6			1				1			
NNS54	ADV	7									1	
NNS01	HI	6										*wa-Ø, o-Ø*
NNS09	HI	6						2				
NNS12	HI	7			1							
NNS14	HI	2				1				4		*CP*
NNS19	HI	6	1								1	
NNS22	HI				5		1					*ga-waX2*
NNS25	HI	8										
NNS30	HI	7	1									
NNS31	HI	7						1				
NNS32	HI	8										
NNS33	HI	7										*o-Ø*
NNS37	HI	4						2			1	*o-o*
NNS38	HI	8										
NNS43	HI	8										

Notes: *RC* stands for relative clause, and *Switch* stands for instances when the picture was described starting with the un-highlighted NP. The numbers in parentheses on the "Passive" and "RC" column indicates the frequencies of incorrect use of passives and relative clauses. *CP* on the "other" column stands for causative passive.

Table 2.5: Raw frequencies of the particles that the NNS participants used to produce non-reversible scrambled sentences in Task 2 (continued)

#	Groups	*o-ga/wa*	*(*)ni-ga*	**ga/wa-o*	**ga/wa-ni*	**ni-o*	**o-ni*	Passive	RC	Switch	other
NNS04	LI					3		(1)			*o-Ø, to-o, to-Ø, misc.*
NNS05	LI			8							
NNS06	LI	8									
NNS07	LI	7								1	
NNS13	LI	6				1	1				
NNS15	LI	6								2	
NNS16	LI	8									
NNS17	LI	1		5	1						*CP*
NNS21	LI	7		1							
NNS23	LI	3		1						4	
NNS28	LI	7	1								
NNS34	LI	5								2	*o-o*
NNS35	LI	7									*Ø-o*
NNS36	LI			8							
NNS39	LI	7					1				
NNS41	LI			3	3				(2)		
NNS49	LI			4	2	1					*ga-Ø*
NNS50	LI	4	1								*o-oX3*
NNS51	LI			2	3	1	2				
NNS55	LI			1	5	2					

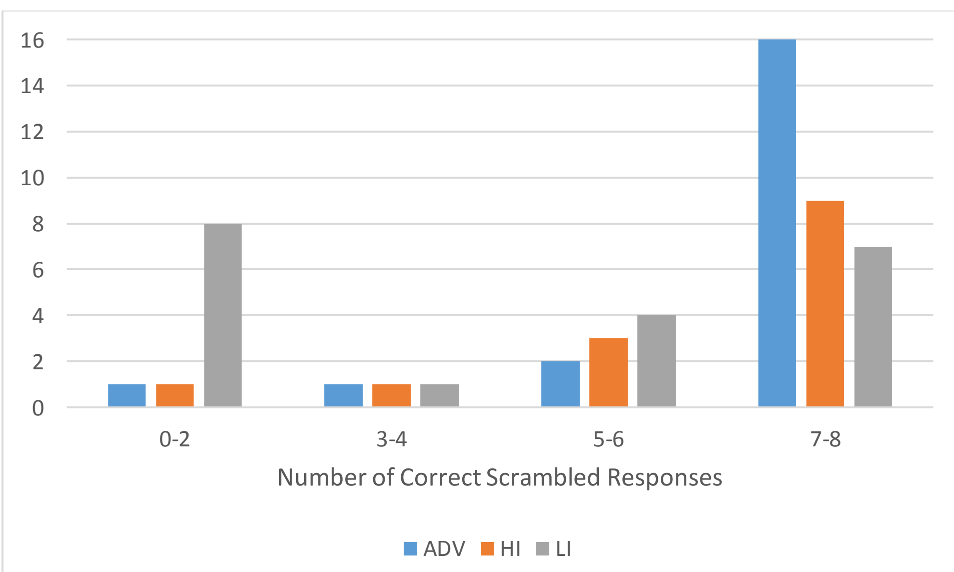


Figure 2.14: The number of participants by proficiency group who produced the non-reversible scrambled sentences correctly in Task 2

Individual differences can be observed in this data as well. While the LI group on average was significantly less accurate in producing non-reversible scrambled sentences, 11 out of 20 LI participants used the correct "*o-ga/wa*" four times (i.e., 50%) or more. On the other hand, 7 LI participants produced sentences in a SOV form 5 times or more when they were prompted to produce non-reversible scrambled sentences.

While the marking of the accusative NP with *–ni* was quite frequent with the reversible scrambled pictures (28 participants marked the accusative NP with *–ni* at least once), *–ni* marking of the accusative NP was far less frequent with the non-reversible scrambled pictures (only three instances by three participants). As with Task 1, this suggests that the participants were more familiar with the non-reversible verbs and that they were more accurate in the choice of the correct particle (i.e., *-o*) with these verbs, although the frequent use of the *–ni* particle with the reversible sentences could also be attributed to the participants' conscious efforts to mark the agent, as suggested above.

When compared with the reversible pictures, the use of the passive was also less frequent with the non-reversible scrambled pictures. While 13 participants used (or attempted to use) the passive to describe the reversible scrambled pictures, there were only three participants who attempted to use the passive with the non-reversible scrambled pictures. The use of inanimate subjects is less common in Japanese sentences, and the passive construction in Japanese often implies that the subject (patient) is adversely affected by the action denoted by the verb. Thus, a passive sentence with an inanimate subject generally has a reduced grammaticality as compared to a passive sentence with an animate subject.[21] The less frequent use of the passive construction

[21] Kuroda (1979) argues that the subject of the *–ni* direct passive has to be an "affectee" by the passive verb *rare*, whereas the subject of the *–ni yotte* passive is not imposed such an restriction. Thus, in the following examples (Kuroda, 1979: 330-331), there is a grammaticality contrast because 'Fermat's theorem' cannot be interpreted as an affectee. Refer to Hoshi (1999) for a theoretical review of Japanese passives.

in the description of the non-reversible scrambled pictures thus seems to suggest that the NNS participants might have been sensitive to this reduced grammaticality of the passive construction with an inanimate subject. Similarly, the marking of the second noun with –*kara* ("from"), which was somewhat frequent with the reversible scrambled pictures, never occurred with the non-reversible counterparts. This is possibly due to the NNS participants' familiarity with the non-reversible verbs as well as their understanding that inanimate subjects are less frequently used in Japanese sentences.

As with the reversible scrambled condition, the raw frequencies of the instances in which the first NP was marked with a subject marker and the second NP was marked with an object marker were counted. The subject marker was defined as –*ga* or –*wa* and the object markers were –*o* or –*ni*. There were 70 instances of wrong NP1-*ga* and 95 instances of wrong NP2-*o*/*ni*. A Chi-square test indicated that the differences in observed frequencies approached significance ($\chi^2(1) = 3.788$, $p = 0.052$), indicating that the participants tended to supply –*o* or –*ni* for the second NPs more frequently than –*ga* for the first NP.

2.8.3.3 NS participants

Tables 2.6 and 2.7 present the raw frequencies of the particles and the sentence types that each NS participant used to describe the reversible and non-reversible scrambled pictures.

---

a. *Fermat-no teiri-ga John-ni syoomeis-are-ta.
-Gen theorem-Nom -by prove-Pass.-Past
'Fermat's theorem$_i$ was affected by John's proving it$_i$.'

b. Fermat-no teiri-ga John-ni yotte syoomeis-are-ta.
-Gen theorem-Nom -to owing prove-Pass.-Past
'Fermat's theorem was proven by John.'

Table 2.6: Raw frequencies of the particles that the NS participants used to produce reversible scrambled sentences in Task 2

#	*o-ga/wa*	**ga/wa-o*	Passive	RC	Ca	Sw	other
NS01	6		1				*to-ga*
NS02	1		7				
NS03	6		2				
NS04			7				*no-ga*
NS05	3		3				*to-ga, morau*
NS06	7	1					
NS07			6		1	1	
NS08	2		6				
NS09			6		1	1	
NS10	3		5				
NS11	1		7				
NS12	3	1	3			1	
NS13	3		5				
NS14	1		5	1			*add*
NS15	7		1				
NS16	4		2		1		*ni-ga*
NS17	2		4	1			*o-o*
NS18	1		7				
NS19	4	2	2				
NS20			5	3			

Notes: *Ca* stands for causative. *Sw* stands for switch, indicating the instances when the picture was described starting with the un-highlighted NP. In the "other" column, *morau* stands for the –temorau construction, and *add* indicates that additional words or expressions were used during the picture description.

As mentioned above, while the NNS participants worked on the fill-in-the-blank task (Task 1) prior to the picture description task, and therefore had recent exposure to the scrambling structure, the NS participants did not have such a prior exposure to scrambling. The higher frequency of passives by the NS participants, especially with the reversible scrambled pictures (19 out of 20 NS participants used passive at least once) seems to reflect the lack of the recent activation of the scrambling structure.

Table 2.7: Raw frequencies of the particles that the NS participants used to produce non-reversible scrambled sentences in Task 2

#	*o-ga/wa*	Passive	RC	Switch	other
NS01	8				
NS02	8				
NS03	7	1			
NS04		8			
NS05	8				
NS06	8				
NS07	4	4			
NS08	6	2			
NS09		8			
NS10	7			1	
NS11	7	1			
NS12	4		3	1	
NS13	8				
NS14	2		6		
NS15	8				
NS16	8				
NS17	4		4		
NS18	1	4	3		
NS19	7				*o-ni*
NS20	2		6		

The frequent use of passives and relative clauses by the NS participants with the non-reversible pictures is of interest. As mentioned above, a passive sentence with an inanimate subject sounds somewhat awkward. Also, none of the NS participants used relative clauses to describe the canonical or the filler pictures. The high frequency of passives and relative clauses (despite the fact that they sound more awkward than scrambled sentences) then seems to reflect a rather low activation level of the scrambling structure among some of the NS participants. Two NS participants (NNS04, 09) never produced a scrambled sentence during the task.

2.8.4 Reaction time data

The reaction time (RT) data were collected to examine if there are any psychological costs associated with the production of scrambled sentences. As discussed above, the RT data were measured manually, and therefore, accuracy to a millisecond level is not assured. The results are only suggestive for this reason.

The RT data were analyzed by participant group. Only correct responses based on the "lax" scoring were included in the RT analysis. The raw RTs were transformed using the reciprocal transformation before they were analyzed with a linear mixed-effects model. Residuals that are 2.5 standard deviations away from the model were removed from analysis. Figure 2.15 shows the trimmed raw reaction times for each sentence type by participant group.

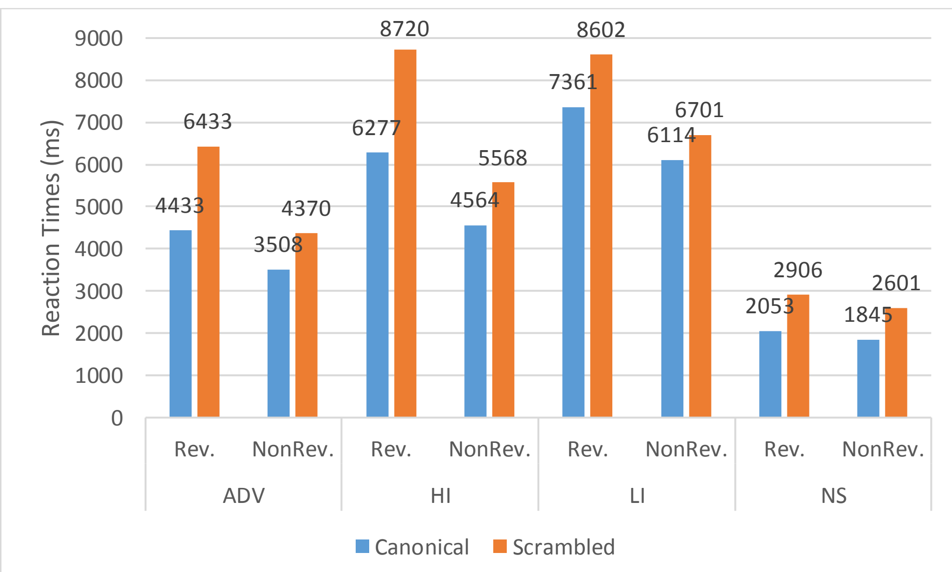


Figure 2.15: Reaction times for each sentence type by participant group in Task 2

The data from the HI group indicated a significant interaction between Reversibility and Word order, and therefore, the following linear mixed-effect model was used for analysis:

```
inverseRT ~ Reversibility + Word order + (1 | subject) + (1 |
```

`item)`. For the other participant groups, the model without the interaction was used to analyze the data. Note that the stimuli for the reversible and non-reversible conditions were strictly separate, and therefore, any effect of reversibility is only suggestive.

For the ADV group, 19 data points out of 755 (2.52%) were trimmed prior to the analysis. The main effects of Reversibility and Word order were significant (Reversibility: $t = 4.962$, $p < 0.0001$; Word order: $t = 5.033$, $p < 0.0001$). The interaction of Reversibility and Word order was not significant ($\chi^2(1) = 1.0935$, $p = 0.2957$).

For the HI group, 12 data points out of 603 (1.99%) were trimmed. The main effects of Reversibility and Word order were significant (Reversibility: $t = 5.788$, $p < 0.0001$; Word order: $t = 3.517$, $p = 0.0005$). The interaction of the two factors was marginally significant ($t = 2.031$, $p = 0.0431$). For the reversible sentences, the difference in RT between the canonical and scrambled sentences was significant ($t = 3.893$, $p = 0.0001$) while the difference was not significant for the non-reversible sentences ($t = 1.237$, $p = 0.2169$).

For the LI group, 16 data points out of 676 (2.37%) were trimmed. The main effects of both Reversibility and Word order were significant (Reversibility: $t = 3.54$, $p = 0.0014$; Word order: $t = 2.755$, $p = 0.0062$). The interaction of the two factors was not significant ($\chi^2(1) = 0.0478$, $p = 0.8269$).

Finally, for the NS participants, 10 data points out of 731 (1.37%) were trimmed. The main effects of Reversibility and Word order were significant (Reversibility: $t = 3.281$, $p = 0.0024$; Word order: $t = 10.142$, $p < 0.0001$). The interaction of the two factors was not significant ($\chi^2(1) = 1.1042$, $p = 0.2933$).

As quite evident from Figure 2.15, the RT data indicated that it took longer for the participants to start producing the scrambled sentences than the canonical sentences, and that it took longer for them to start producing the reversible sentences than the non-reversible sentences, although the difference may be attributed to the different items in the two conditions. A significant interaction of Reversibility and Word order was observed only in the HI group.[22]

2.8.5 Discussion

The analysis of the mean accuracy rate data based on the "lax" scoring indicated different effects of reversibility on scrambling among the proficiency groups. For the reversible sentences, no significant difference was observed in the interaction of Reversibility and Word order among the proficiency groups, suggesting that all three of the NNS groups were equally less accurate in producing the scrambled sentences as compared to the canonical sentences. For the non-reversible sentences, on the other hand, while all three of the groups were also less accurate in producing the scrambled sentences, the LI group was significantly less accurate in their production of the scrambled sentences as compared to the other groups. Although the interaction of Reversibility and Word order among the NNS groups were not significant for the reversible sentences, the accuracy rates of the reversible scrambled sentences were significantly different among the groups; the ADV group indicated higher accuracy and the LI group lower. Thus, as for the relationship of L2 learners' general proficiency and their production of scrambled sentences, it can be said that there is a positive relationship.

While a positive relationship between L2 learners' proficiency and their production of scrambled sentences was observed, the raw frequency data also indicated individual differences within the participant groups. For instance, while the accuracy rate of the non-reversible

[22] The difference in RT between the reversible canonical and scrambled sentences was larger for the HI group than the LI group, suggesting that the overuse of the canonical template may come from experience with the L2 rather than transfer from L1 or a universal perceptual strategy.

scrambled sentences was significantly lower among the LI participants, 11 out of 20 LI participants produced such sentences using the correct "*o-ga/wa*" particles 50% or more of the time. On the other hand, there were 7 participants in the same group who produced sentences in an SOV form 5 times or more in describing the non-reversible scrambled pictures.

The mean accuracy rates for the reversible scrambled sentences were consistently lower than those of the non-reversible scrambled sentences. This probably was partly due to the NNS participants' relative unfamiliarity with the reversible verbs used in the task. The data based on the "strict" scoring indicated that the accuracy rates in producing the reversible canonical sentences among the HI and LI participants were less than 50%, showing that the participants might not have known the correct use of the reversible verbs in many instances. However, the lower accuracy rates for the reversible scrambled sentences may also be due to the application of the NNV strategy. There were a total of 39 NNS participants out of 54 (72%) who produced sentences with an SOV form at least once in their description of the reversible scrambled pictures, while there were a total of 15 participants who did the same with the non-reversible scrambled sentences. Also, there were 95 total instances of the "*ga/wa-o/ni*" and "*ga-wa*" (i.e., SOV) forms in the description of the reversible scrambled pictures whereas there were 60 such instances with the non-reversible scrambled pictures. While the marking of the accusative NP with *–ni* could have been due to the participants' conscious efforts to form passives as discussed above, the frequent occurrences of the SOV forms during the description of the reversible scrambled pictures suggest a more frequent manifestation of the NNV strategy in producing reversible scrambled sentences.

The application of the NNV strategy may also be manifested as the frequent occurrences of "NP1-*ga/wa*" during the description of the reversible scrambled pictures. There were 125 instances of "NP-*ga*/*wa*" for the reversible scrambled pictures while there were 70 instances for the non-reversible scrambled sentences. The highlighted words in the reversible scrambled pictures were animate (humans), which conforms to the SOV template in Japanese, and therefore, the participants might have been more inclined to attach –*ga* after the first NPs.

The RT data indicated that not only the NNS participants but the NS participants took longer to start producing scrambled sentences, suggesting an additional psychological cost. The frequent use of passives among the NS participants suggested a rather low activation level of the scrambling structure among some of them. Two NS participants never produced a scrambled sentence during the task.

2.9 Grammatical knowledge and production performance

2.9.1 Mean response accuracy data

For the purpose of examining the effect of the different tasks (fill-in-the-blank and picture description), the mean accuracy rates of the two tasks were compared by proficiency groups using the linear mixed-effects modeling. The accuracy rates used in this analysis were the ones based on the "lax" scoring. The fixed effects were Reversibility (reversible, non-reversible), Word order (canonical, scrambled), and Task (Task 1, Task 2). Random effects were Subject and Verb. Verb represents the actual verbs used in the sentence items. Verb was entered as a random effect instead of Item because the sentence items between the two tasks were not exactly identical but the same verbs were used for the two tasks. For the data of the ADV and HI groups, the interactions of the three fixed effects were significant, and therefore, the following model was used for analysis: `Accuracy ~ Reversibility * Word order * Task + (1 | subject) + (1 | verb)`. On the other hand, there was no significant interactions among factors for the data of the LI group, and thus the following model was used for analysis:

```
Accuracy ~ Reversibility + Word order + Task + (1 | subject) +
(1 | verb).
```

The mean accuracy data for the ADV group presented in Figures 2.1 and 2.6 are repeated in Figure 2.16.

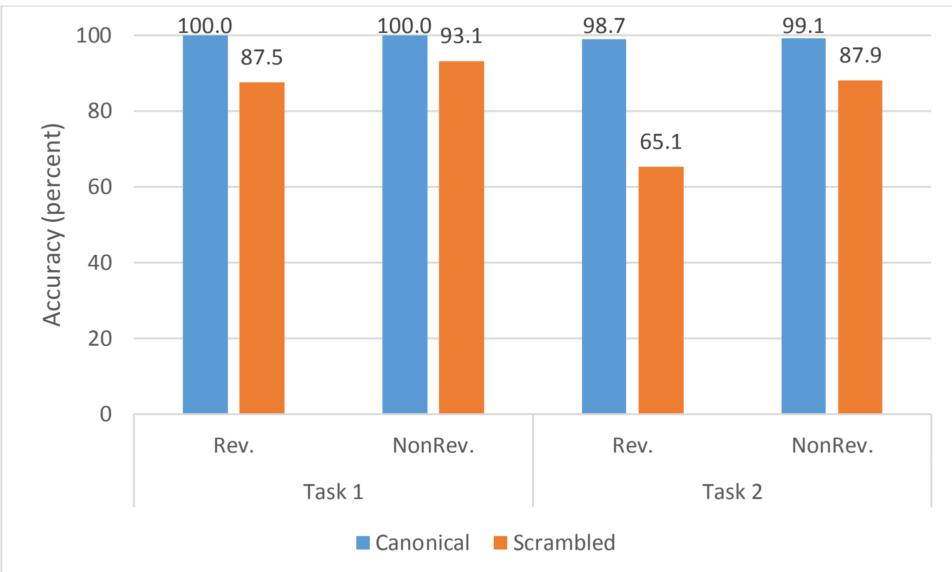


Figure 2.16: Mean accuracy rates of Tasks 1 and 2 for the ADV group

The main effect of Task was significant (t = 5.629, p < 0.0001). The main effects of Reversibility and Word order were also significant (Reversibility: t = 4.791, p < 0.0001; Word order: t = 12.321, p < 0.0001). The interaction of Word order and Task as well as that of Reversibility and Task were significant (Word order-Task: $\chi^2(1)$ = 25.596, p < 0.0001; Reversibility-Task: $\chi^2(1)$ = 7.4995, p < 0.0062). The interaction of Reversibility and Word order was also significant ($\chi^2(1)$ = 24.272, p < 0.0001).

The pairwise comparisons of the sentence types between the tasks indicated that the difference between the two tasks was significant for the reversible scrambled sentences (t =

8.765, $p < 0.0001$) as well as for the non-reversible scrambled sentences ($t = 2.557$, $p = 0.0107$). Differences for either of the canonical sentence types were not significant.

The pairwise comparisons of interactions between the tasks indicated that the interaction of Task and Reversibility was significant for the scrambled sentences ($t = 4.581$, $p < 0.0001$), but it was not significant for the canonical sentences ($t = 0.101$, $p = 0.9197$). The interaction of Task and Word order was significant for the reversible sentences ($t = 6.085$, $p < 0.0001$), but it was not significant for the non-reversible sentences ($t = 1.569$, $p = 0.1169$). The three-way interaction was significant ($t = 3.264$, $p = 0.0011$). Thus, the result indicated that the ADV participants were less accurate for both types of scrambled sentences in Task 2, but were less accurate in producing the reversible scrambled sentences than the non-reversible scrambled sentences.

The mean accuracy data of the two tasks by the HI group are repeated in Figure 2.17.

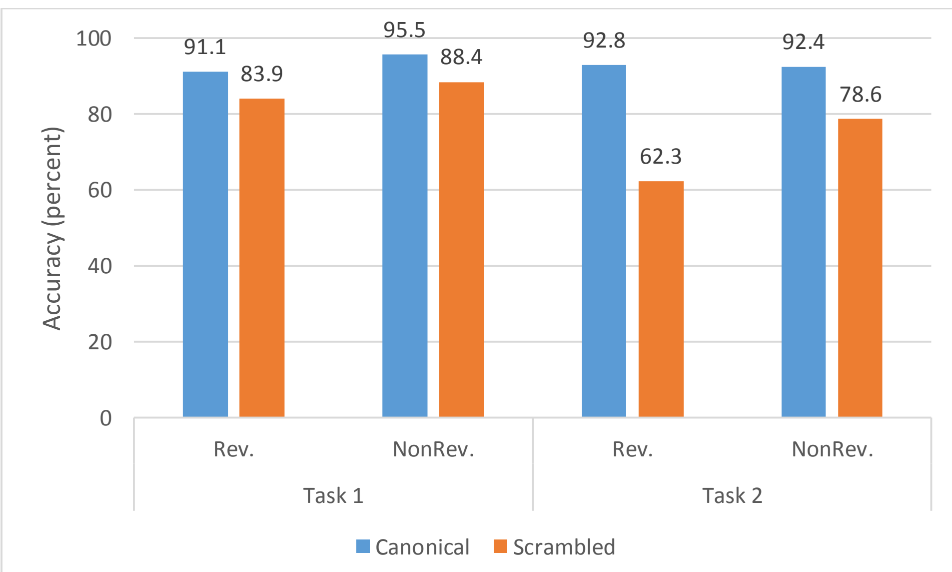


Figure 2.17: Mean accuracy rates of Tasks 1 and 2 for the HI group

The main effect of Task was significant ($t = 3.738$, $p = 0.0002$). The main effects of Reversibility and Word order were significant (Reversibility: $t = 2.824$, $p = 0.0049$; Word order: $t = 6.727$, $p < 0.0001$). The interaction of Task and Word order was significant ($\chi^2(1) = 13.189$, $p = 0.0005$), but that of Task and Reversibility was not significant ($\chi^2(1) = 0.567$, $p = 0.4514$). The interaction of Reversibility and Word order only approached significance ($\chi^2(1) = 3.5885$, $p = 0.0581$).

The pairwise comparisons of the sentence types between tasks indicated that the difference between the two tasks was significant for the reversible scrambled sentences ($t = 5.038$, $p < 0.0001$) as well as for the non-reversible scrambled sentences ($t = 2.273$, $p = 0.0233$). The differences for the canonical sentence types were not significant.

The pairwise comparisons of interactions between tasks indicated that the interaction of Task and Reversibility was marginally significant for the scrambled sentences ($t = 1.978$, $p = 0.0482$), but it was not significant for the canonical sentences ($t = 0.813$, $p = 0.4163$). The interaction of Task and Word order was significant for the reversible sentences ($t = 3.887$, $p = 0.0001$), but it was not significant for the non-reversible sentences ($t = 1.099$, $p = 0.2723$). The three-way interaction was marginally significant ($t = 1.982$, $p = 0.0478$). The result therefore indicated that, similarly to the ADV group, the HI participants were particularly less accurate in producing the reversible scrambled sentences in Task 2.

The mean accuracy data of the two tasks by the LI group is repeated in Figure 2.18.

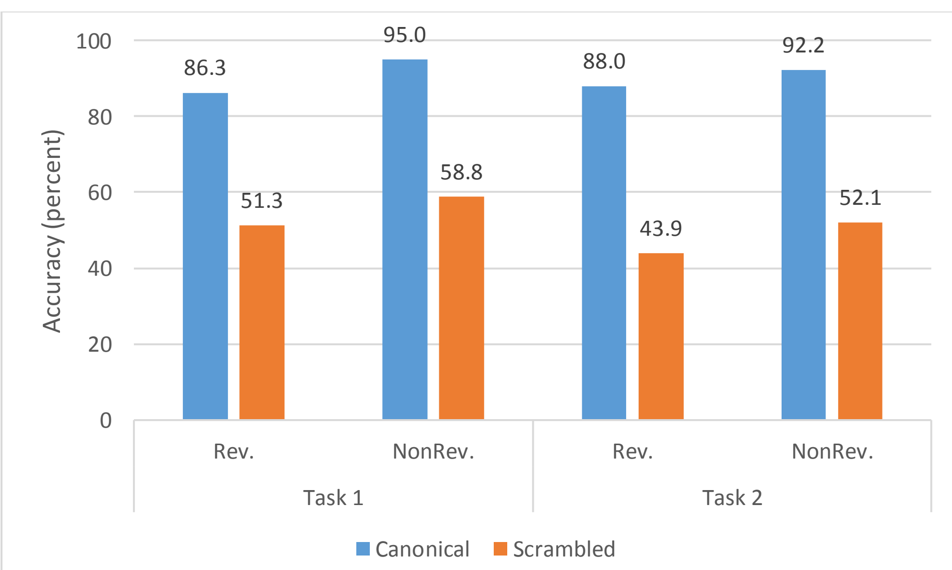


Figure 2.18: Mean accuracy rates of Tasks 1 and 2 for the LI group

The main effect of Task only approached significance (t = 1.917, p = 0.0555). The main effects of Reversibility and Word order were significant (Reversibility: t = 3.727, p = 0.0002; Word order: t = 19.964, p < 0.0001). No interactions were significant. The result therefore indicated that the mean accuracy rates in the two tasks were generally similar for the LI participants.

Overall, the quantitative comparisons of the accuracy rates in the two tasks indicated that the participants were less accurate in their production (Task 2) of the scrambled sentences than in Task 1. The ADV and HI groups were particularly less accurate in their production of the reversible scrambled sentences although their production accuracies of the same sentence type were still significantly higher than that of the LI group.

2.9.2 Raw frequency data

For the purpose of examining a) whether the errors in Task 2 may be attributed to the overuse of the SOV template and b) whether there is a correspondence between grammatical knowledge and production performance on the individual level, the raw frequency data of the two tasks were compared.

2.9.2.1 Reversible sentences

Table 2.8 shows a) the number of participants who provided erroneous SOV-based responses at least once, b) the number of erroneous SOV-based responses, and c) the number of erroneous "NP1-subject" marking and "NP2-object" marking, in the reversible scrambled condition of the two tasks.[23] The "SOV-based" form is defined as "NP1-*ga* or –*wa* AND NP2-particle other than –*ga* or –*wa* ". The "*ga-wa*" responses were also considered as an SOV-based form.

[23] For the frequency of SOV-based responses in Task 2, only the final responses were included. That is, if a participant self-corrected and produced an OSV form as the final response, the initial error was not counted. For the

Table 2.8: The number of participants who provided SOV-based responses and the frequencies of the SOV-based forms in the reversible scrambled condition in Tasks 1 and 2

	No. of Participants	Freq. of SOV form	Freq. of NP1-subj.	Freq. of NP2-obj.
Task 1	29	68	69	110
Task 2	39	110	125	119

There were more participants who used the SOV forms erroneously in Task 2. There were also more instances of erroneous SOV-based responses and the "NP1-*ga*/*wa*" markings in Task 2. As discussed above, the high frequency of "*ga-ni*" marking in the production of reversible scrambled sentences may have been partly due to the participants' conscious efforts to mark the agent with *–ni*. However, the overall increase of the SOV forms in Task 2 seems to indicate that the errors during the production of reversible scrambled sentences were due (at least partly) to the overuse of the SOV template.

Table 2.9 presents the same data as Table 2.8 but the data are shown by participant group. Because the SOV-based errors were quite frequent among the LI participants in both tasks and also because there were a large number of responses in non-target forms (e.g., passives) in Task 2, simple comparison of the participant groups is not possible. However, the data in Table 2.9 seem to indicate a substantial increase in responses based on the SOV form among the participants in the ADV and HI groups (especially in the latter group), suggesting a frequent application of the SOV template during sentence production.

frequencies of erroneous "NP1-subject" and "NP2-object" markings, on the other hand, correct marking of the NP preceded by an error was given the point of 0.5.

Table 2.9: The number of participants who provided SOV-based responses and the frequencies of the SOV-based forms in the reversible scrambled condition in Tasks 1 and 2 (by group)

Groups	Tasks	No. of Participants	Freq. of SOV	Freq. NP1-subj.	Freq. NP2-obj.
ADV	1	9	14	14	19
	2	13	26	33.5	31.5
HI	1	4	7	7	17
	2	10	24	26.5	22.5
LI	1	16	47	48	74
	2	16	60	65	65

Figures 2.19, 2.20, and 2.21 present numbers of participants in each group who made a certain numbers of erroneous SOV-based responses, "NP1-subject" marking, and "NP2-object" marking in the two tasks.[24] As can be observed, the number of participants who made no SOV-based errors (as well as NP1-subject and NP-2 object markings) decreased in Task 2 as compared to Task 1, while the number of participants who made one or two SOV-based errors increased Task 2, especially among the ADV and HI group.

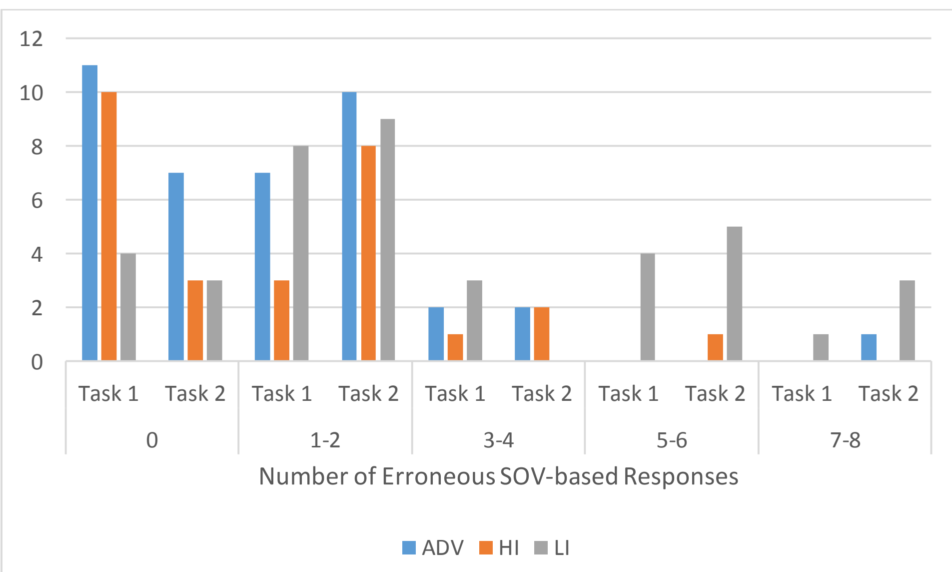


Figure 2.19: The number of participants by group who made a certain numbers of erroneous SOV-based responses in the reversible scrambled condition in Tasks 1 and 2

[24] For the Task 2 data in Figures 2.20 and 2.21, the number of errors by each participant was round up. That is, if a participant had the errors of 1.5 times, it was listed as "2" in the two figures.

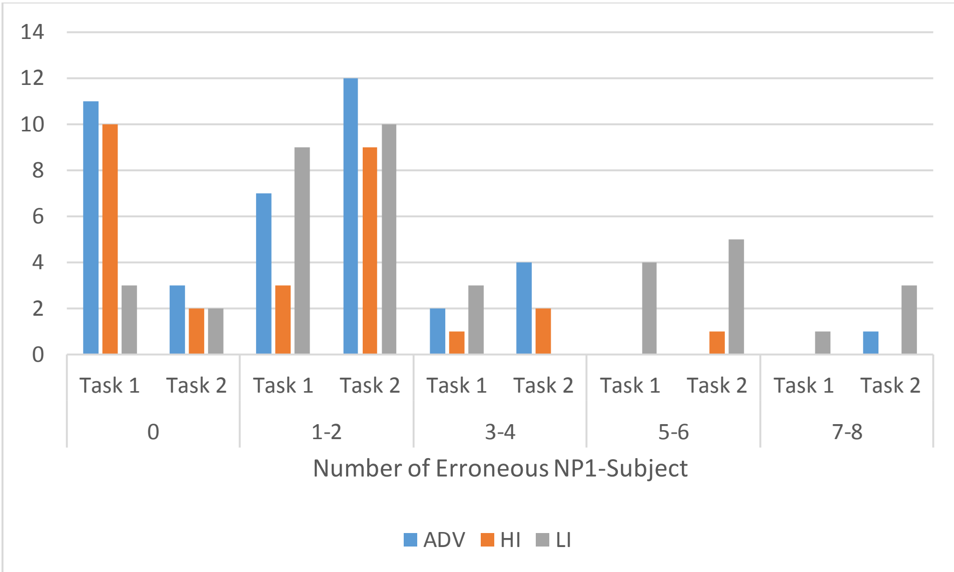


Figure 2.20: The number of participants by group who made a certain numbers of erroneous NP1-subject responses in the reversible scrambled condition in Tasks 1 and 2

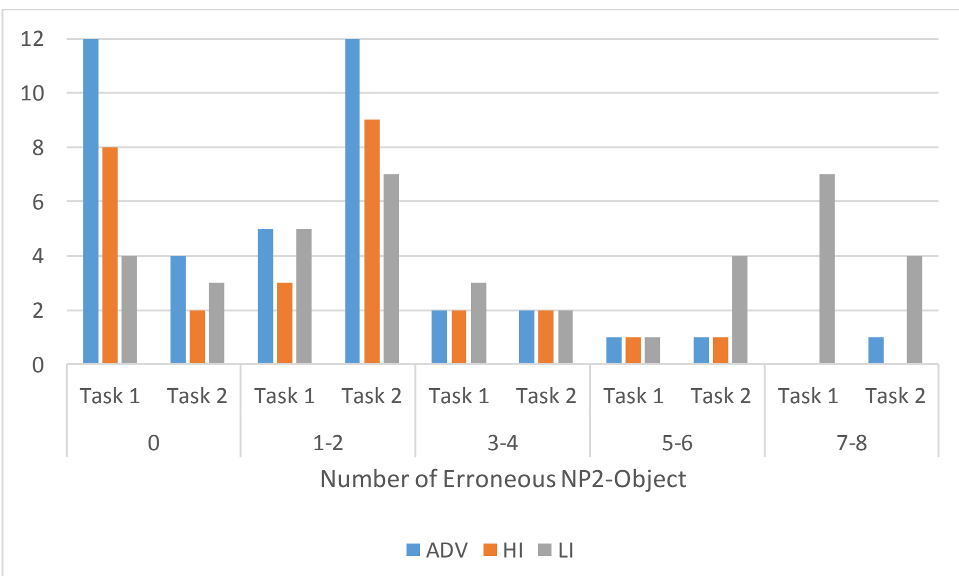


Figure 2.21: The number of participants by group who made a certain numbers of erroneous NP2-object responses in the reversible scrambled condition in Tasks 1 and 2

Table 2.10 shows the ID numbers of the participants who completed or produced the reversible scrambled sentences four times or fewer during the two tasks. Table 2.11 is a cross table showing, by task, the number of participants in terms of the frequency of OSV sentence completion/production.

Table 2.10: The list of participants who completed or produced reversible scrambled sentences equal to or less than 50% of the time in Tasks 1 and 2

Task 1	ADV	18, 44
	HI	22, 31, 37
	LI	04, 05, 17, 35, 36, 39, 41, 49, 51, 55
Task 2	ADV	03, 10, 18, 44, 46, 53
	HI	01, 09, 19, 31, 33
	LI	04, 05, 13, 17, 21, 23, 35, 36, 41, 49, 51

Table 2.11: The number of participants in terms of the frequency of the OSV sentence completion/production in Tasks 1 and 2 (reversible scrambled condition)

		Task 2	
		≥ 5	≤ 4
Task 1	≥ 5	a) 29	b) 11
	≤ 4	c) 3	d) 11

The numbers of participants in cells a) and d) in Table 2.11 indicate a positive correspondence in the participants' performances between the two tasks. The 29 participants (out of 40) who completed the reversible scrambled sentences correctly five times or more in Task 1 also produced sentences of the same type correctly five times or more in Task 2. The 11 participants (out of 14) who completed the reversible scrambled sentences only four times or fewer in Task 1 also produced such sentences correctly only four times or fewer in Task 2. It is speculated that these 11 participants did not know that the OSV word order is permissible in Japanese and/or did not know how to formulate OSV sentences.

On the other hand, knowing the grammar did not seem to ensure the correct production of reversible scrambled sentences. There were 11 participants (NNS03, 10, 46, 53 in ADV; NNS01, 09, 19, 33 in HI; NNS13, 21, 23 in LI) who demonstrated the grammatical understanding but produced correct reversible OSV sentences only 50% of the time or less. Moreover, when the number of correct responses in Task 2 was subtracted from the number of correct responses in Task 1 for each of the 11 participants, there were seven participants whose difference was 3 or 4. As it was possible that some participants experienced difficulties interpreting the pictures, the responses of the seven participants with the reversible canonical pictures were counted. Based on the "lax" scoring, there were three participants among the seven who made errors in describing the reversible canonical pictures: NNS01 four times, NNS 09 twice, and NNS 33 once. Thus, perhaps with the exception of NNS01, at least six participants seem to have experienced greater difficulty in producing reversible scrambled sentences despite their grammatical knowledge of scrambling.[25]

There were three participants (NNS22, 37, and 55) who completed the reversible scrambled sentences correctly four times or fewer but produced sentences of the same type correctly five times or more. Closer inspection of their responses showed that the first two participants provided the correct responses four times in Task 1 and five times in Task 2. Therefore there was no large discrepancy between their grammatical knowledge and their production of reversible scrambled sentences by these participants. NNS55 provided only two correct responses in Task 1 but was correct five times in Task 2. However, this participant used the "*ni-ga*" particles four times to describe the reversible canonical pictures, and therefore, the seemingly correct descriptions of the scrambled pictures may be attributed to the overuse of the "*ni-ga*" response pattern.

---

[25] A majority of the 11 participants still produced correct reversible scrambled sentences four times in Task 2.

2.9.2.2 Non-reversible sentences

Table 2.12 shows a) the number of participants who provided erroneous SOV-based responses at least once, b) the number of erroneous SOV-based responses, and c) the number of erroneous "NP1-subject marker" and "NP2-object marker", in the non-reversible scrambled condition of the two tasks.

Table 2.12: Frequencies of the SOV-based forms in the non-reversible scrambled condition in Tasks 1 and 2

	No. of Participants	Freq. of SOV form	Freq. of NP1-subj.	Freq. of NP2-obj.
Task 1	15	59	60	83
Task 2	15	62	70	95

The numbers of the participants who used an erroneous SOV form in one of the two tasks were the same (15 participants) while 12 of them made such responses in both tasks. There were more instances of erroneous "NP1-*ga*/*wa*" and "NP2-*o*/*ni*" responses in Task 2, but the increase of the responses based on the SOV form was only 3. That there were three participants (NNS53 in ADV and NNS 12 and 14 in HI) who never used an SOV form in the non-reversible scrambled condition of Task 1 but used such a form in Task 2 (once for each participant) might be taken as evidence that these participants used the SOV template in producing non-reversible sentences. However, given that the increase of SOV-based forms in Task 2 was minimal, it appears that errors in Task 2 were not due to the overuse of the SOV template during production. Rather, the errors in production seem to have been directly linked to a lack of grammatical knowledge.

Table 2.13 shows the ID numbers of the participants who completed or produced the non-reversible scrambled sentences four times or fewer during the two tasks. Table 2.14 is a cross table showing, by task, the number of participants in terms of the frequency of the OSV sentence completion (Task 1) and production (Task 2).

Table 2.13: The list of participants who completed or produced non-reversible scrambled sentences equal to or less than 50% of the time in Tasks 1 and 2

Task 1	ADV	18, 44
	HI	22, 37
	LI	04, 05, 17, 36, 39, 41, 49, 51, 55
Task 2	ADV	18, 44
	HI	22, 37
	LI	04, 05, 17, 23, 36, 41, 49, 51, 55

Table 2.14: The number of participants in terms of the frequency of the OSV sentence completion/production in Tasks 1 and 2 (non-reversible scrambled condition)

		Task 2	
		≥ 5	≤ 4
Task 1	≥ 5	a) 41	b) 1
	≤ 4	c) 1	d) 11

As observed in Table 2.14, there is a very close relationship between the participants' performances on the two tasks. The 41 participants out of 42 who completed the non-reversible scrambled sentences correctly five times or more in Task 1 also produced sentences of the same type correctly five times or more in Task 2. The 11 participants out of 12 who completed the non-reversible scrambled sentences only four times or fewer in Task 1 also produced such sentences correctly only four times or fewer in Task 2. Nine of the participants in cell d) of Table 2.14 were among the eleven participants in cell d) of Table 2.11. These nine participants, therefore, appear to have experienced difficulty with the scrambling structure regardless of the sentence type (reversible and non-reversible).

The one participant in cell b) of Table 2.14 (NNS23) correctly described the non-reversible scrambled pictures three times, but there were also four instances of "switching" in which the participant started the description of the pictures with the animate noun first although the inanimate noun was highlighted in the picture. Therefore, the higher error rate in the production may be attributed to the high frequency of switching. The one participant in cell c) (NNS39) actually completed the non-reversible scrambled sentences correctly four times. Thus, it can be said that individual participants' grammatical knowledge overall corresponded very well with their production performance for the non-reversible scrambled sentences.

Taken together, the comparison of the participants' performance between the two tasks indicated that the relation between grammatical knowledge of scrambling and production was somewhat different between reversible and non-reversible sentences. For the reversible sentences, while the participants' performance in the two tasks mostly corresponded, there were cases where participants who demonstrated a grammatical understanding of scrambling did not necessarily perform well in production. There were more errors in producing scrambled sentences especially among the ADV and HI groups, and at least some of the errors seem to have been caused by the overuse of the SOV template although there is a possibility that the participants' relative unfamiliarity with reversible verbs might have contributed to the errors. For the non-reversible sentences, on the other hand, the grammatical knowledge of scrambling and production matched very closely, and there was little evidence that the errors in production of scrambled sentences were caused by the overuse of the SOV template.

For the purpose of comparing the reversible and non-reversible conditions further in terms of the participants' performance in the two tasks, correlation coefficients of the two tasks were calculated by condition. The scores used here were based on the "lax" scoring, but unintended correct responses (e.g., correct use of passives) in Task 2 were considered correct (instead of treating them as missing) in this scoring. For the reversible OSV sentences, Pearson $r = 0.789$ (n = 54, $p < 0.001$). For the non-reversible OSV sentences, Pearson $r = 0.916$ (n = 54, $p < 0.001$). This indicates that there was a significant correlation between the participants' grammatical knowledge and their production performance of OSV sentences for both conditions, but the correlation was more robust for the non-reversible sentences.

2.10 General discussion

The two tasks in this chapter attempted to assess L2 Japanese learners' grammatical knowledge and their production of scrambled sentences. While there were substantial individual differences in participants' knowledge of scrambling, even within proficiency groups, the results of Task 1 (fill-in-the-blank task) indicated that the ADV and HI groups performed significantly better than the LI group in completing the scrambled sentences, suggesting that, on average, there is a positive relation between L2 learners' overall proficiency and their grammatical knowledge of scrambling.

There were 13 participants whose accuracy in completing the scrambled sentences in Task 1 was equal to or below 50%. This is of concern given that all participants in this study but one (i.e., NNS22, who actually was among the 13 participants who had difficulty in completing scrambled sentences) completed at least four semesters of classroom instruction in Japanese. It is speculated that one reason for a lack of grammatical knowledge may be the lack of explicit instruction on scrambling. In the short debriefing interview, when the participants were asked if they knew that the OSV word order is permissible in Japanese, 33 participants replied "Yes" while 18 replied "No". Among the participants who replied "No", two said they knew about the flexibility in word order but did not know that the OSV order was allowed. Also, when the participants were asked if they learned explicitly about the word order flexibility in the language classroom, only 20 participants responded that they learned or they "think" they learned about scrambling explicitly. As observed, the use of the OSV word order is rather rare in Japanese, and therefore it would probably be difficult for L2 learners to learn about scrambling word order from natural input alone (especially in written text, which are generally edited). For this reason, it might be beneficial to integrate scrambling into a grammar lesson and practice it in class. As production of scrambling necessarily requires that learners be aware of the function of postpositional particles, such practice may also contribute to the acquisition of the particles.

The results of Task 2 (picture description) indicated that, while the accuracy rates of the scrambled sentences were significantly lower than those of the canonical sentences across the NNS groups consistently, the ADV and HI participants produced scrambled sentences significantly more accurately than the LI participants, indicating a positive relation of L2 learners' proficiency in Japanese and their production of scrambled sentences. Furthermore, the comparison of the results of the two tasks indicated that grammatical knowledge and production performance matched very closely for the non-reversible scrambled sentences. Because the scrambling in simple transitive sentences is rather a simple operation, it is speculated that the participants were generally able to use their grammatical knowledge for the production of non-reversible scrambled sentences in which the animacy contrast of the two nouns clearly signaled which noun was supposed to be the subject/object.

For the reversible scrambled sentences, on the other hand, while grammatical knowledge and production performance generally matched, there appeared to be some discrepancy between the two. There were at least 6 participants who demonstrated grammatical understanding in Task 1 but had some difficulty in producing the reversible scrambled sentences in Task 2. The number of erroneous responses based on the SOV forms indicated an increase of such errors in Task 2, suggesting that the production errors of the reversible scrambled sentences were at least partly due to the overuse of the SOV template. In the reversible scrambled pictures, the highlighted nouns were humans. Given that Japanese, similarly to English, generally prefers an animate subject to inanimate one, the participants might have been more prone to errors based on the SOV template for the reversible scrambled sentences.[26]

Comparison of the two tasks also indicated that, as observed in some of the previous studies, the overuse of the SOV template in production seemed to persist even at a higher proficiency level. The ADV and HI participants, who generally demonstrated a good grammatical understanding of scrambling, still seemed to experience some difficulty in producing the reversible scrambled sentences possibly due to interference by the overuse of the canonical template. As reviewed above, Otsu (1994a) demonstrated that the gap between grammatical knowledge and performance during the comprehension task among L1 Japanese children could be alleviated by providing a natural context before the scrambled sentences. It would be very interesting to observe whether such a method may also be effective in reducing the knowledge/performance gap in L2 production of scrambling.

The NNS participants in the present study were all native speakers of English, and there was evidence that some of them overused the SOV template in completing the tasks. However, as Iwasaki (2003) pointed out, it would be too simplistic to conclude that the overuse of the SOV template is due to a negative transfer from the participants' L1. As Slobin and Bever (1982) demonstrated, children cross-linguistically seem develop and use a canonical sentence template that conforms to the typical features (e.g., word order, inflection, and prosody) of the input

[26] While the results of the present study suggest that the production of reversible scrambled sentences is more difficult than non-reversible scrambled sentences, this point requires further examination. First, the NNS participants in this study seemed to be less familiar with the reversible verbs in the material. Second, because there were two persons in the reversible scrambled pictures, the identification of agent/patient in the pictures might have been intrinsically more difficult for those pictures. Third, the reversible scrambled pictures could be described using the passive construction instead of scrambling, and the description with passive would sound more natural than that with scrambling in many cases. Therefore, it is possible that the competition between the passive and scrambling structures might have caused difficulty during the production of the reversible scrambled sentences.

language. Some of the studies with Japanese children also found that they seem to go through a developmental stage during which they overuse the SOV word order, which is certainly more typical than the OSV order (e.g., Hayashibe, 1975). Moreover, although very infrequent, the NS participants made errors based on the SOV forms in describing the reversible scrambled pictures. The NS participants also indicated a preference for the SOV order over OSV (e.g., description of the non-reversible scrambled pictures with passives). Thus, the NNS participants' overuse of the SOV template may not be the result of L1 interference. Rather, it may be due to a universal perceptual/productive strategy and/or a strategy shaped by the typical features of Japanese sentences (i.e., SOV word order, animate subject) through their acquisition of the target language.

CHAPTER 3: COMPREHENSION OF SCRAMBLED SENTENCES BY L2 LEARNERS OF JAPANESE

3.1 Introduction

This chapter reports the results of the pilot study (sentence correctness decision task) and Task 3 (self-paced reading task) which were aimed at investigating L2 learners' comprehension of scrambled sentences. Online studies that have examined the comprehension of scrambled sentences by NSs of Japanese have provided evidence that there is a psychological cost in processing scrambled sentences. That is, it takes longer to read and comprehend scrambled sentences than canonical sentences. On the other hand, studies that have examined the online comprehension process of scrambled sentences by L2 learners of Japanese are scarce, and it is not entirely clear whether such slowdowns are also experienced by L2 learners when they read and comprehend scrambled sentences. Not only may investigations into the topic deepen our understanding of how scrambled sentences are processed by L2 learners, they may also provide insight as to whether L2 sentence processing is qualitatively different from L1 processing, which has recently been an area of growing interest.

3.2 The Japanese parser and scrambling

Studies in English sentence processing have suggested that the information carried by verbs plays a significant role in the parsing of English (e.g., Trueswell, Tanenhaus, & Garnsey, 1994). Japanese is a left-branching, verb-final language with scrambling and the extensive use of empty categories (Berwick & Fong, 1995). If the Japanese parser relies on verb information as the

English parser does, the parsing of a Japanese sentence will be highly ambiguous until the verb is encountered.[27]

Although various models have been proposed for the processing of Japanese, they can be roughly divided into two types. The first type is the incremental model, which hypothesizes that input is processed incrementally without delay. The obvious advantage of the model is the lower processing load because only one syntactic structure is pursued at a time. The disadvantage of such a parser, on the other hand, is that there is a high probability that the parsing decision may

---

[27] As observed in the following examples, it is sometimes difficult to tell how a sentence will unfold in Japanese, without any contextual clue:

a. John-ga Mary-o…
-Nom -Acc

b. John-ga Mary-o mita.
-Nom -Acc saw
'John saw Mary.'

c. John-ga Mary-o mita toki…
-Nom -Acc saw when
'When John saw Mary…'

d. John-ga [e_i Mary-o mita] gakusei$_i$-o…
-Nom -Acc saw student-Acc
'John… (did something)… to the student who saw Mary.'

e. Ø [John-ga Mary-o mita]-no-o shitteiru.
-Nom -Acc saw Nominalizer-Acc know
'I know that John saw Mary.'

f. John-ga Mary$_i$-o Bill-ni e_i shookaishita
-Nom -Acc -Dat introduced
'John introduced Mary to Bill.'

As a sentence starts with two NPs as in (a), it can be a simple clause such as (b). When (b) is followed by a complementizer such as *toki* ('when') as in (c), the clause is interpreted as a subordinate clause. Alternatively, if (b) is followed by an NP as in (d), *Mary-o mita* is understood as a relative clause that modifies the NP. (Japanese does not have relative pronouns such as *who* and *which* in English or a special verbal ending that signals a relative clause as in Korean. Therefore, that a clause is a relative clause is not noticed until the head noun is encountered, although prosody would help the disambiguation in speech.) Japanese employs empty categories (*pro*) extensively, especially when the entities that the empty categories specify are known or assumed from the context. In (e), *John-ga Mary-o mita* is followed by the nominalizer *–no* and the accusative case marker, and therefore, the clause is understood as the direct object of the sentence. Although the agent of *shitteiru* ('know') is not explicitly stated in (e), it is assumed that the agent is the speaker ('I'). In addition to the extensive use of empty categories, Japanese allows scrambling. As (f) shows, the *John-ga Mary-o* order may be derived from a ditransitive sentence whose canonical word order is NP-ga NP-ni NP-o.

be incorrect. One influential model proposed to mitigate the disadvantage of the incremental parser is the "Ranked Flagged Information Paced Parser" by Inoue and Fodor (1995). Within this model, while the parser attaches incoming information incrementally onto the syntactic tree during the first-pass reading, it places a "flag" where an alternative parsing possibility exists. If the first-pass parsing turns out to be incorrect, the parser goes back to the flag and pursues another structure. Hence, the parser can avoid the cost of analyzing the syntactic structure of the sentence all over again.

The second type of parsing model is the delay model, which hypothesizes that parsing decisions are delayed until they can be made safely. Within this model, while the probability of misanalysis is reduced, the psychological processing load may increase because unanalyzed items need to be held onto until they can be attached to the syntactic tree. One example of the delay model is the head-driven model by Pritchett (1991, 1992). The head-driven parser delays parsing decisions until the syntactic head is encountered. According to the model, no syntactic tree is built until the verb becomes available.

The two models make different predictions in parsing canonical sentences and their scrambled counterparts (Mazuka, Itoh, & Kondo, 2002).

(3.1) a. [$_{IP}$ John-ga [$_{VP}$ ringo-o tabeta]]
-Nom apple-Acc ate
'John ate an apple.'

b. [$_{IP}$ Ringo$_{i}$-o [$_{IP}$ John-ga [$_{VP}$ t_{i} tabeta]]]
Apple-Acc -Nom ate

As observed in Chapter 1, it is considered, from the generative perspective, that a scrambled sentence such as (3.1b) is derived from the canonical counterpart (3.1a) (e.g., Saito & Hoji, 1983). For the scrambling operation, an additional IP node is adjoined to the syntactic tree, and the accusative NP is moved to the higher [Spec IP] position, leaving a trace. If the parsing of Japanese sentences is strictly incremental, when the first argument of (3.1b) *ringo* (apple) is received, a slight delay in processing may be observed because the parser may sense that the sentence being parsed is a scrambled one, and it is necessary to assume a trace (gap) and build an additional IP node. Another distinct processing delay is expected when the second argument, *John*, is received because it is now apparent to the parser that the structure being parsed is scrambled. The incremental model does not expect a processing slowdown in (3.1b) at the verb. The syntactic tree having been built, the verb simply needs to be filled into the respective slot.

On the other hand, the head-driven model does not expect a processing difference either at the first or second argument because no syntactic decisions are made before the verb becomes available. However, a significant processing difference is expected at the verb. Because the scrambling structure requires another IP node and processing of a gap, a delay in reading time is expected at the verb for scrambled sentences.

Alternatively, if we assume that Japanese has a non-configurational "flat" structure (e.g., Farmer, 1984; Hale, 1980), no processing differences between (3.1a) and (3.1b) are expected at any points in the two sentences.

3.3 Literature review

3.3.1 Psycholinguistic investigations on comprehension of scrambled sentences by NSs of Japanese

One of the first psycholinguistic studies that investigated the processing of scrambled sentences in Japanese was conducted by Nakayama (1995). As observed above, a scrambled element is considered to leave a trace when it is moved higher up in the syntactic tree. In this study, the researcher attempted to examine the psychological reality of traces left by scrambling.

The participants read sentences presented on a computer screen one phrase at a time, at their own pace. Each sentence was followed by a probe word, and the participants were asked to decide whether or not the probe word appeared in the sentence they had just read. If there is a psychological reality of traces motivated by scrambling, the reaction time for the probe words should be shorter for the scrambled sentences. That would be so because the probe words are expected to be activated at the trace positions located before the verb.

Nakayama set four conditions listed below (3.2). The sentences (3.2a) and (3.2c) are both canonical sentences, but the former has the probe word in the subject modifier position while the latter in the object modifier position. Likewise, (3.2b) and (3.2d) are scrambled sentences, but (3.2b) has the probe word in the subject modifier position and (3.2d) in the object modifier position. The underlining indicates the positions of the probe words.

(3.2) a. <u>modifier</u> NP-ga modifier NP-o V.
b. modifier NP-o <u>modifier</u> NP-ga V.
c. modifier NP-ga <u>modifier</u> NP-o V.
d. <u>modifier</u> NP-o modifier NP-ga V.

The results indicated that (3.2a) evoked significantly slower reaction times than (3.2b), while (3.2c) evoked significantly faster reaction times than (3.2d). That is, when the probe words appeared later in the sentence, the reaction times were faster. However, the reaction times to the probes in the scrambled sentences were not different from those in the canonical sentences. Hence, the experiment did not detect a scrambling effect in terms of the reactivation of traces left by scrambling.

Although the previous studies on passives in English (Bever & McElree, 1988) and on unergative and unaccusative sentences in Japanese (Nakayama, 1990, 1991; cited in Nakayama, 1995) detected the psychological reality of the traces, Nakayama (1995) on scrambled sentences did not. Nakayama speculates that this may be due to the different status of traces in the previous studies and the current study. The traces in the previous studies (passive and unaccusative constructions) are obligatory, while all the scrambled stimuli sentences in the 1995 study are grammatical without scrambling, suggesting that the traces in the sentences "may be optional (and absent)" (Nakayama, 1995: 270).[28]

While Nakayama (1995) attempted to observe the reactivation of traces left by scrambling, and therefore tried to observe the scrambling effect somewhat indirectly, Yamashita (1997) directly examined whether any extra psychological load is involved in the processing of scrambled sentences. In the first experiment of the study, the researcher prepared four types of ditransitive sentences as shown below. (3.3a) is the canonical sentence. In (3.3b), the indirect object is moved to the sentence-initial position, while in (3.3c), the direct object is scrambled. In (3.3d), both the indirect and direct objects are scrambled. The sentences were presented using a self-paced moving-window technique.

(3.3) a. NP-ga NP-ni NP-o V.
b. NP-ni NP-ga NP-o V.
c. NP-o NP-ga NP-ni V.
d. NP-ni NP-o NP-ga V.

The result showed no significant differences in reading time across the conditions at any position, suggesting that there was no increased difficulty in processing scrambled sentences.

[28] Miyamoto and Takahashi (2004), on the other hand, found a scrambling effect in terms of the speed of probe recognition, using sentence stimuli based on Nakayama's (1995).

While the studies reviewed above did not observe an extra processing load for scrambled sentences, other studies have found evidence of additional psychological cost involved in the processing of scrambled sentences. Mazuka, Itoh, and Kondo (2002) used eye tracking and self-paced reading techniques to examine the processing of scrambled sentences. Their stimuli included the following simple and complex canonical/scrambled sentences.[29]

(3.4) a. *Canonical simple sentence:* [NP-ga NP-o V]
b. *Scrambled simple sentence:* [NP-o NP-ga V]
c. *Canonical sentence with a center embedding:* [NP-ga [modifier phrase] NP-o V]
d. *Scrambled sentence with a center embedding:* [NP-o [modifier phrase] NP-ga V]

The results of their eye-movement experiment indicated, for overall reading time, that (3.4d) was read significantly more slowly than (3.4c), indicating a processing cost associated with scrambled sentences. On the other hand, differences in overall reading time between (3.4a) and (3.4b) did not reach a level of statistical significance. A more detailed analyses of the eye-tracking data indicated that the readers tended to gaze longer and made more regressive eye movements at the second argument position in the scrambled sentences, while the accusative-marked NP that was fronted in the scrambled sentence did not take longer to read.

Mazuka et al.'s self-paced reading experiment showed a similar trend although there was a difference between simple and complex sentences. For complex sentences (3.4c & d), the reading times for the canonical and scrambled sentences were not significantly different from each other at the first NP position, but they were significantly different at the second NP position. On the other hand, for the simple sentences, the reading times for the canonical and

[29] Mazuka et al.'s (2002) experiments included another sentence type in which the modifier clause was placed at the beginning of the sentence. This was to test the suggestion that one of the motivations for the scrambling structure in Japanese is to avoid center embedding. The results indicated that processing cost was indeed reduced when the sentence did not have center embedding.

scrambled sentences were not significantly different either at the first or second NP position, which suggests that the self-paced reading method might not have been sensitive enough to detect slowdowns in very simple scrambled sentences such as (3.4b). Overall, however, Mazuka et al.'s data indicated that slowdowns in comprehending scrambled sentences take place at the second NP before the verb is encountered, a result most consistent with the incremental processing model.

Based on their results, Mazuka et al. (2002) suggest that scrambled sentences are parsed in the following way: when the parser encounters an accusative-marked NP in the first argument position, a significant slowdown does not take place because ellipses occur quite frequently in Japanese and it is possible that the sentence being parsed is canonical with the subject omitted. However, when the parser encounters a nominative NP at the second argument position, the parser identifies the sentence as scrambled, and it needs to create an additional IP node and to process the gap. It is the identification of scrambling and the reanalysis of the sentence structure that is realized as longer reading times.

Although Mazuka et al.'s results indicated the additional cost in processing scrambled sentences, it is not clear what the cause of the increased processing load is. Is it because an additional IP node needs to be created in scrambled sentences, or because the gap caused by scrambling needs to be identified and processed, or both? Miyamoto and Takahashi (2002) used sentences with VP-internal scrambling of dative and accusative arguments to answer this question. The researchers tested the following two types of sentences. Because scrambling takes place within the VP in their design, if there is a difference in reading time in the two sentences, the difference can be attributed to the processing of the gap rather than the need to create another IP node.

(3.5) a. *Canonical:*
[Subject [$_{RC}$ ec$_j$ NP-Dat NP-Acc ditransitive-verb]
NP$_j$-Acc verb-Comp] NP-Nom report-V.

b. *Scrambled:*
[Subject [$_{RC}$ ec$_j$ NP$_i$-Acc NP-Dat gap$_i$ ditransitive-verb]
NP$_j$-Acc V-Comp] NP-Nom report-V.

(Miyamoto & Takahashi, 2002: 170)

The results of the self-paced reading experiment indicated that, although comprehension accuracy was low for both canonical and scrambled sentences (61% and 65%, respectively), the scrambled sentences were read significantly more slowly than the canonical sentences at the argument prior to the ditransitive verb (the underlined NPs in the example above). No significant differences in reading speed were observed at any other regions. Thus the results suggested that slowdowns in scrambled sentences might be attributed to the processing of the gap but not necessarily to the need to create an additional clause boundary. Because there was a significant slowdown in the scrambled sentences before the verb but no slowdown was observed at the verb, the result of the study is also consistent with the incremental view of the Japanese parser.

Muraoka, Tamaoka, and Miyaoka (2004) investigated whether there would be a processing difference when the scrambled NP is accusative or dative. The conditions were the following:

(3.6) a. NP-ga NP-o V. (accusative, canonical)
b. NP-o NP-ga V. (accusative, scrambled)
c. NP-ga NP-ni V. (dative, canonical)
d. NP-ni NP-ga V. (dative, scrambled)

In this study, the entire sentence was projected on a computer monitor, and the participants made judgments about whether each sentence was correct as quickly and accurately as possible (sentence correctness decision task). The results showed that the canonical sentences were read significantly faster than the scrambled sentences, but there was no significant difference in reading speed between the accusative and dative conditions. Based on the results, the researchers concluded that the parser starts to identify gap positions motivated by scrambling, regardless of the types of case markers. Other studies that employed the sentence correctness decision task have also found evidence of additional psychological cost during the comprehension of scrambled sentences (Chujo, 1983; Tamaoka, Sakai, Kawahara, Miyaoka, Lim, & Koizumi, 2005).

As observed, psycholinguistic studies on Japanese scrambling have provided growing evidence of additional psychological loads in processing scrambled sentences.[30] Such additional costs seem to be experienced before the verb is encountered, which suggests that sentence processing in Japanese is incremental and the parser starts to build the syntactic tree before the information provided by the verb becomes available. The eye-tracking and self-paced reading experiments in Mazuka et al. (2002) indicated that the locus of slowdown during the comprehension of OSV sentences by NS readers is the second NP where it becomes apparent to the parser that the sentence being parsed is scrambled.

3.3.2 L2 studies on comprehension of scrambled sentences

To the best of the author's knowledge, aside from studies that investigated the comprehension of scrambled sentences from the perspective of the competition model (some of which were reviewed in the previous chapter), there have been only a few psycholinguistic studies on L2 comprehension of scrambled sentences in Japanese. However, the results of those studies appear conflicting.

[30] Other methods that have been employed to investigate Japanese scrambling include cross modal priming technique (Nakano, Felser, & Clahsen, 2002), ERP (Ueno & Kluender, 2003), and fMRI (Kim et al., 2009), all of which have observed scrambling effects.

Hara (2009) investigated the processing of two types of VP-internal scrambling (scrambling of "IO-DO" order to "DO-IO" order) by advanced learners (L1 Chinese, Korean) of Japanese and by NSs as a control group. The conditions were the following:

(3.7) a. *Canonical-ordered condition:*
NP-wa [locative phrase]-de NP-ni NP-o V.
-Top -Loc -Dat -Acc

b. *Short-scrambling condition:*
NP-wa [locative phrase]-de NP-o$_i$ NP-ni t_i V.

c. *Long-scrambling condition:*
NP-wa NP-o$_i$ [locative phrase]-de NP-ni t_i V.

It was found, through a moving-window self-paced reading experiment, that L1 Korean learners slowed down at the gap-implicating positions for the short-scrambling condition (no intervening phrase between the scrambled accusative and the dative arguments), while they did not show such slowdowns for the long-scrambling condition (intervening adjunct phrase between the scrambled accusative and the dative arguments). NSs, on the other hand, slowed down in the long-scrambling sentences but not in the short-scrambling sentences. L1 Chinese participants did not indicate reading slowdowns in either of the scrambling types. Hara interpreted the result as evidence that Korean participants, whose L1 shows scrambling properties similar to Japanese, were capable of processing the syntactic gap created by scrambling. In the processing of long-scrambling, however, Korean participants might not have engaged in syntactic gap processing because the task at hand placed a high demand (longer distance between the filler and the gap) on their computational resources.

Mitsugi and MacWhinney (2010) also examined the processing of scrambled sentences by adult L2 learners of Japanese (L1 Chinese, English, Korean) and native speakers. The processing of three types of scrambled sentences along with canonical sentences was tested:

(3.8) a. *Canonical:*
NP-de NP-ga NP-ni NP-o V.
-Loc -Nom -Dat -Acc

b. *Dative scrambling:*
NP-de NP-ni$_i$ NP-ga t_i NP-o V.

c. *Accusative scrambling:*
NP-de NP-o$_i$ NP-ga NP-ni t_i V.

d. *Dative-accusative scrambling:*
NP-de NP-ni$_i$ NP-o$_j$ NP-ga t_i t_j V.

Contrary to Hara (2009), the results of the self-paced reading experiment indicated that there was no significant difference in reading time among the different sentence types, except that the L1 Korean participants read dative-accusative sentences significantly more slowly than the accusative scrambled sentences at the first NP position.

Tamaoka (2005) tested the comprehension of SOV and OSV sentences by L1 Chinese learners of Japanese. The method employed was a sentence correctness decision task, in which the entire sentence was presented on a computer monitor, and the participants were asked to make a decision as to whether the sentence is correct as quickly and accurately as possible. The results indicated that OSV sentences were read and comprehended significantly more slowly than SOV sentences and that OSV sentences elicited more decision errors than SOV sentences. Based on the results, Tamaoka suggests that Chinese learners of Japanese in the study were capable of processing the gap motivated by scrambling in a manner similar to native speakers.

3.3.3 L2 sentence processing and the Shallow Structure Hypothesis

Whether the L2 parser is qualitatively different from the L1 parser has been a matter of debate in the L2 sentence processing literature. It appears that the issue has been discussed even more frequently since the publication of Clahsen and Felser (2006). In that article, the authors reviewed a number of psycholinguistic studies on morphology and sentence processing among adult NSs, NS children, and late L2 learners. Based on their analyses, the authors suggested that, while L2 learners' morphological processing is comparable to NS adults and children, their sentence processing mechanisms are fundamentally different from those of NS adults and children. That is, L2 syntactic representation for online comprehension is "shallower and less detailed than those of native speakers" (p.32), and therefore, while NSs compute complex hierarchical structures and abstract elements during sentence comprehension, such computation may be absent in the sentence comprehension of adult L2 learners; L2 sentence processing seems to rely on lexical-semantic and pragmatic information instead (the Shallow Structure Hypothesis: SSH).

Clahsen and Felser (2006) argue that there is an interpretive ambiguity in the findings of previous L2 studies on filler-gap dependencies in English. Juffs and Harrington (1995), for instance, found slowdown at the post matrix verb position when their L2 learner participants read sentences such as *Who did Anne say her friend likes __?* (p. 505). Clahsen and Felser maintain that it is ambiguous as to whether such slowdown, which is taken to be evidence of online filler integration, is the result of a gap-filling strategy based on structural information or of a more "shallow" strategy based on the verb's subcategorization information. The distinction is difficult to make because in English the verb is placed earlier in the sentence, and the structure-based gap-filling and the verb-driven processing may take place at the same post-verbal position.

SSH is partly based on the result of Marinis, Roberts, Felser, and Clahsen (2005), which investigated whether L2 learners of English would reactivate the antecedent at the intermediate gap position created by long-distance *wh*-extraction. The target stimulus Marinis et al. employed was the following (p. 61):

(3.9) The nurse *who*$_i$ the doctor argued e'_i that the rude patient had angered e_i is refusing to work late.

The rationale is that, if the reader reactivates the antecedent at the intermediate gap (e'_i), the complementizer *that* would trigger such reactivation, and therefore, longer reading time is expected at this position. At the same time, the activation of filler at the intermediate position would predict facilitation of filler integration. Thus, shorter reading times would be expected at the region of the base gap (e_i). The results indicated that this was exactly the case for NSs; their reading speed slowed down at *that* but they read *had angered* faster, compared with the condition which did not have the intermediate gap.[31] On the other hand, none of the L2 groups (L1 Chinese, German, Greek, and Japanese) showed such a slowdown at *that*, although they slowed down at the *had angered* region, suggesting filler integration at this position. Given the results, Marinis et al. suggest that the L2 filler-gap processing is driven by lexical information rather than grammatical requirements.

Although Marinis et al.'s sentence stimuli allowed them to distinguish verb-driven integration from structure-based gap filling, Hara (2009) points out that Marinis et al.'s results may not necessarily indicate that L2 learners do not engage in structure-based gap-filling. In Marinis et al., the L2 participants were assigned to comprehend relatively long sentences with two-step *wh*-

[31] The stimuli without the intermediate gap was like the following: The nurse *who*$_i$ the doctor's argument about the rude patient had angered e_i is refusing to work late (Marinis et al., 2005: 61).

movement. Therefore, it is possible that participants in Marinis et al. were unable to engage in syntactic gap-filling due to the complexity of the materials, which might have placed a high demand on their computational resources.

As Hara (2009) points out, verb-final languages such as Japanese can serve as a proper testing tool to examine whether L2 sentence processing is qualitatively different from L1 sentence processing. Because of the verb-finality, the distinction between verb-based and structure-based gap-filling should be relatively clear. If L2 sentence processing does involve structure-based gap-filling, slowdown in reading time will be observed at a position before the verb. If, instead, the gap integration is verb-driven, slowdown will be observed at the verb.

While SSH maintains that L2 sentence processing may be qualitatively different from the L1 processing, studies that examined the relationship between proficiency in L2 and L2 sentence processing have provided evidence that L2 sentence processing at a very high proficiency level may not be qualitatively different from its L1 counterpart. For instance, Steinhauer, White, Cornell, Genesse, and White (2006) reports that "high proficiency" (as measured by a cloze test score) late learners of English (L1 Chinese and French) elicited a similar ERP pattern (LAN/P600, which are believed to reflect syntactic processing) as NSs, while "low proficiency" learners exhibited a different pattern.[32] Therefore, the finding suggests that late L2 learners, regardless of typological difference between L1 and L2, are able to attain NS-like syntactic processing at a high proficiency level.

Hopp (2006) investigated the processing of subject/object ambiguity caused by scrambling in German. The participants in the study were L1 English and Dutch learners of German who were divided into two groups based on a C-test score ("advanced" and "near-native"). The stimuli

[32] LAN effect was absent in low proficiency learners, which indicated that those learners did not show automatic detection of violations.

mainly included the two types: disambiguation by case marker and disambiguation by verbal agreement. Below is an example of agreement-disambiguation sentence (Hopp, 2006: 372):

(3.10)	Er	sagt,	dass	die	$Kellnerin_{SG}$	die	$Kollegen_{PL}$	gesehen	hat_{SG}.
	He	says	that	$the_{NOM/ACC}$	waitress	$the_{NOM/ACC}$	colleagues	seen	has

‘He says that the waitress saw the colleagues.’

While subject/object disambiguation takes places at the NP in the case-marker-disambiguation sentences, disambiguation does not take place until the clause-final auxiliary verb in the agreement-disambiguation sentences. The results of the self-paced reading experiment indicated that, for NSs and near-natives, OS word order caused slowdown at the NPs for the case-disambiguation sentences, while slowdown took place at the disambiguating verb for the agreement-disambiguation sentences. The advanced group, on the other hand, only exhibited general slowdown at the end of the sentence, but there was no indication of slowdown at the disambiguating regions. Hopp interprets the finding as the evidence that “endstate non-native processing and native processing are qualitatively identical” (p.369). Native-like processing among very proficient learners is reported also in Frenck-Mestre (2002).

3.4 The present study

The primary purpose of the present study is to investigate how scrambled sentences are read and comprehended by L2 learners of Japanese. In the following, the results of two experimental tasks, the pilot study and Task 3, will be reported. The pilot study used the sentence correctness decision task, in which the entire sentence was presented on the computer monitor and the participants were asked to make a judgment as to whether the sentence was plausible as quickly and accurately as possible. The sentence correctness decision task might not be as informative as eye-tracking or self-paced reading methods because it does not indicate where the slowdowns take place in scrambled sentences (Miyamoto, 2008). However, previous studies (e.g., Chujo, 1983; Tamaoka, 2005) have observed the scrambling effect using this method, and therefore, it was felt that the method was suitable for a preliminary investigation of the processing of scrambled sentences by L2 learners. The word-by-word online processing of scrambled sentences was examined in Task 3, which employed a self-paced reading method.

As with the other tasks in the study, both reversible and non-reversible sentences were included in the pilot study and in Task 3. In reversible sentences, because both nominative and accusative NPs are animate, the only information signaling that the sentence is scrambled are the case markers. In non-reversible sentences, on the other hand, the "inanimate-animate" NP order would also signal that the sentence is scrambled because it is less likely, from our general knowledge, that an inanimate subject would do something to an animate object. If we assume that scrambling causes slowdowns and more difficulty in comprehension, the following two predictions may be made about L2 learners' use of animacy and case-marker information: a) If L2 learners integrate the animacy information of NPs in their comprehension of scrambling, there will be different degrees of scrambling effects between reversible and non-reversible sentences, and b) If L2 learners integrate case-marker information in their comprehension of scrambling, longer reading time is expected for reversible scrambled sentences as compared to reversible canonical sentences.

The specific research questions addressed in the present study are the following:

a) Do L2 learners of Japanese experience slowdowns and greater difficulty in comprehending scrambled sentences as compared to canonical sentences? If they do slow down, where is the locus of the slowdown?

b) Is there a relationship between L2 learner's general proficiency in Japanese and how they process scrambled sentences?

c) Are there qualitative differences between L1 and L2 processing of scrambled sentences?

d) What are the effects of reversibility on L2 learners' processing of OSV sentences? Do L2 learners integrate the animacy of NPs during their processing of the scrambling structure?

3.5 Method for the pilot study[33]

3.5.1 Participants

Twenty-four people (12 females and 12 males) in the University of Arizona community who learned Japanese as an adult (18 years of age or older) participated in the pilot study. They were all native speakers of English, except two native speakers of Chinese (one female and one male). None of the NNS participants in the pilot study were among those who took part in the tasks in the present study (Tasks 1-6). They varied in terms of ages and experiences with the Japanese language. Ages ranged from 20 years to 54 years, but the majority of participants (19 of 24) were in their 20s. The length of their Japanese studies ranged from 1.5 to 20 years, with the mean of 4.3 years and the median of 3 years. Thirteen participants had stayed in Japan for more than a month.

In addition to the NNS individuals, the 20 NS participants who took part in the other tasks in the present study also participated in the pilot study.

[33] The data from the pilot study were analyzed using *t*-tests in Shigenaga (2012). See the article for method details and complete material of the pilot study.

3.5.2 Materials and design

There were four types of sentence stimuli as shown below. Two counter-balanced lists of items were created which included eight sentences in each of the four sentence types. Thirty-two sentences to elicit “no” responses were included, for a total of 64 sentences in each item list.

(3.11) a. *Reversible canonical:*
Taro-ga Kazuko-o mita.
-Nom -Acc saw
‘Taro saw Kazuko.’

b. *Reversible scrambled:*
Kazuko-o Taro-ga mita.
-Acc -Nom saw

c. *Non-reversible canonical:*
Taro-ga shatu-o kita.
-Nom shirt-Acc wore
‘Taro wore a shirt.’

d. *Non-reversible scrambled:*
Shatu-o Taro-ga kita.
Shirt-Acc -Nom wore

3.5.3 Procedure

The participants were randomly assigned to one of the two item lists, which included eight practice items followed by the 64 sentences. The sentence items were presented with the DMDX software (Forster & Forster, 2003) on a laptop computer. Each task began with the appearance of a string of asterisks ‘**********’ in the middle of the monitor for 1000 milliseconds (ms), followed by the sentence item which appeared in the middle of the monitor for 10 seconds or until the correctness decision was made. The participants were instructed to decide, as quickly and accurately as possible, whether the sentence was plausible by pressing the designated keys.

The presentation of stimuli was randomized, so that no participant would see the stimuli in the same order.

3.6 Results of the pilot study

The RT and accuracy data were analyzed using linear mixed-effects modeling in R. The data from the NNS and NS participants were analyzed separately. For both the RT and accuracy data, the fixed effects were Reversibility and Word order, and the random effects were Subjects and Items. A mixed-effects model with the interaction of Reversibility and Word order was applied to the datasets first. As will be observed, however, none of the datasets indicated a significant interaction of the fixed effects, and thus, the results below are based on the analyses using a model without the interaction term. For the purpose of observing the effect of Word order for the reversible and non-reversible sentences, the results based on the model with the interaction will also be reported. The results using a random slopes model will be reported if the model was justified by the data and if the results are substantially different from those without random slopes.

Trials in which an error was made were omitted from the analysis of RT. The raw RT data were transformed with a reciprocal transformation before they were analyzed using the linear mixed-effects modeling. Residuals that were 2.5 standard deviations away from the linear mixed-effects model were removed from analysis.

3.6.1 Reaction times

Figure 3.1 presents the mean reaction times (RT) for each sentence type by the NNS and NS groups.

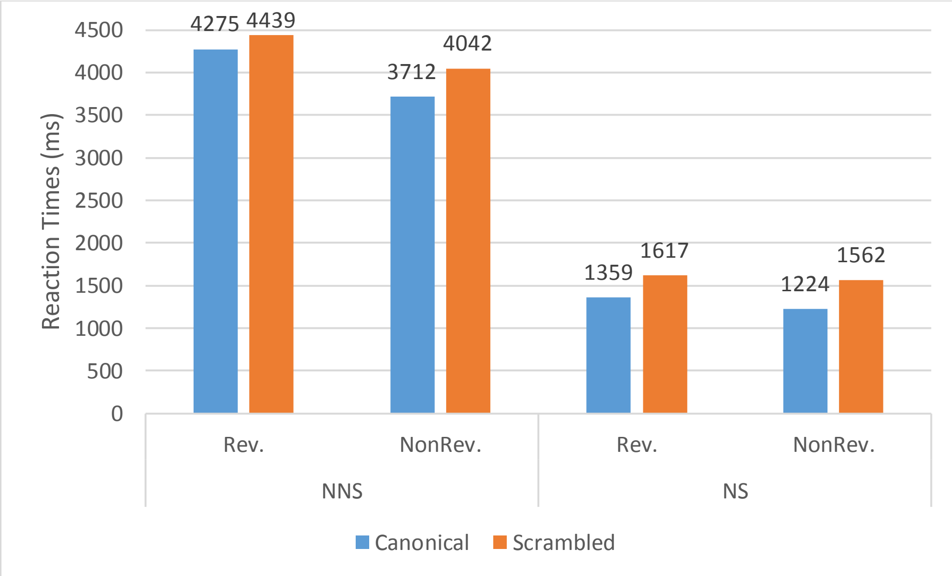


Figure 3.1: Mean reaction times for each sentence type by the NNS and NS groups (the pilot study)

For the NNS group, 9 data points out of 692 (1.3%) were trimmed. The main effects of Reversibility and Word order were significant (Reversibility: t =3.701, p = 0.0009; Word order: t =3.546, p = 0.0005). The interaction of Reversibility and Word order was not significant ($\chi^2(1)$ = 1.4728, p = 0.2249). On the other hand, a mixed-effects model with the interaction of Reversibility and Word order indicated that the effect of Word order was not significant for the reversible sentences (t = 1.592, p = 0.1129), while it was significant for the non-reversible sentences (t = 3.396, p = 0.0008).

For the NS group, 10 data points out of 629 (1.59%) were trimmed. The main effect of Word order was significant (t = 8.71, p < 0.0001), while the main effect of Reversibility was not significant (t = 1.511, p = 0.141). The interaction of the two fixed effects was significant (t = 2.341, p = 0.0198). The pairwise comparison indicated that the effect of Word order was significant for both the reversible (t = 4.584, p < 0.0001) and non-reversible sentences (t = 7.803, p < 0.0001). With an analysis using a random slopes model (by-subject for the effect of

Reversibility), the interaction of the two fixed effects did not reach the level of significance (t = 1.821, p = 0.0695).

3.6.2 Accuracy rates

Figure 3.2 shows the mean accuracy rates for each sentence type.

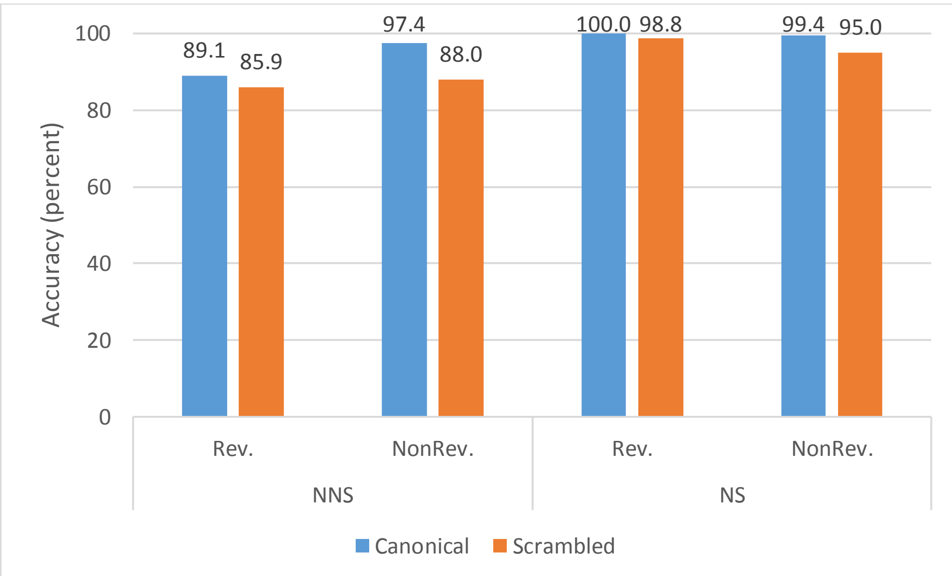


Figure 3.2: Mean accuracy rates for each sentence type by the NNS and NS groups (the pilot study)

For the NNS group, the main effects of both Reversibility and Word order were significant (Reversibility: t = 2.402, p = 0.0227; Word order: t = 2.986, p = 0.0029). The interaction of the fixed effects was not significant ($\chi^2(1)$ = 2.2341, p = 0.135). With a random slopes model (by-subject for the effect of Reversibility), the main effect of Reversibility was not significant (t = 1.788, p = 0.0869). A mixed-effects model with the interaction of Reversibility and Word order indicated that the effect of Word order was not significant for the reversible sentences (t = 1.057, p = 0.2911), while it was significant for the non-reversible sentences (t = 3.17, p = 0.0016).

For the NS group, the main effects of both Reversibility and Word order were significant (Reversibility: $t = 2.144$, $p = 0.0324$; Word order: $t = 2.757$, $p = 0.006$). The interaction did not reach the level of significance ($\chi^2(1) = 2.3611$, $p = 0.1244$). A mixed-effects model with the interaction term indicated that the effect of Word order was not significant for the reversible sentences ($t = 0.867$, $p = 0.386$), while it was significant for the non-reversible sentences ($t = 3.036$, $p = 0.0025$).

3.6.3 Discussion

The result of the pilot study indicated that the scrambled sentences caused longer reaction times and lower accuracy rates among the NS and NNS participants, suggesting that L2 learners of Japanese experience additional psychological cost in reading and comprehending scrambled sentences as do NSs. The result also indicated that the scrambling effect on RT and accuracy rates was more robust in the non-reversible sentences than the reversible ones for both participant groups. This seems to suggest that the NNS participants in the pilot study integrated the animacy information of the NPs in a manner similar to NSs when they read the scrambled sentences.

On the other hand, the use of the case-marker information by the NNS participants in the pilot study might have been different from that of the NSs. While it took significantly longer for the NSs to read and comprehend the reversible scrambled sentences than the canonical ones, the difference did not reach the level of significance for the NNS group. Due to the way in which the task was set up, it was possible for the participants to judge a reversible sentence correct even when they failed to identify which of the two NPs was the subject/object. (“The girl saw the boy.” vs. “The boy saw the girl.”) It is thus quite probable that some NNS participants overlooked the case markers and did not notice scrambling in some of the reversible sentences. In those instances, reanalysis of the sentence structure (i.e., addition of an IP node and processing of the gap) could not have taken place, which might have resulted in the non-significant scrambling effect in the reversible sentences.[34] On the other hand, the observed significant difference in the NS group as well as in previous studies (e.g., Muraoka et al., 2004) suggests that the NS participants integrated the case-marker information more consistently, thereby correctly identifying scrambling in the reversible sentences.

3.7 Method for Task 3

A self-paced reading task was used in Task 3 for the purpose of identifying the locus of slowdown when L2 learners of Japanese read scrambled sentences.

3.7.1 Participants

The 54 NNS and 20 NS participants in other tasks in the present study participated in Task 3.

3.7.2 Materials and design

There were four types of sentence stimuli as shown below. (See Appendix 3A for a list of target items in Task 3.) A locative phrase (e.g., ookii ginkoo-de, 'at the large bank') was placed between the nominative and accusative NPs since the results of Mazuka et al. (2002) suggested that a self-paced reading method might not be sensitive enough to detect slowdowns in very simple OSV sentences.

[34] It is also possible that the simplicity of the stimuli sentences might have reduced observable effects among the NNS participants in the pilot study (cf. Mazuka et al., 2002).

(3.12) a. *Reversible canonical:*

Nichiyoobi-ni Tatsuya-ga ookii ginkoo-de okaasan-o sanjuppun
Sunday-on -Nom large bank-in mother-Acc thirty-minutes
matta rashii.
waited it seems
'It seems that Tatsuya, in the large bank, waited for his mother for thirty minutes on Sunday.'

b. *Reversible scrambled:*

Nichiyoobi-ni okaasan-o ookii ginkoo-de Tatsuya-ga sanjuppun
Sunday-on mother-Acc large bank-in -Nom thirty-minutes
matta rashii.
waited it seems

c. *Non-reversible canonical:*

Senshuu isha-ga atarashii ofisu-de meeru-o yukkuri kaita youda.
last week doctor-Nom new office-in email-Acc slowly wrote it appears
'It appears that the doctor, in the new office, wrote the email slowly last week.'

d. *Non-reversible scrambled:*

Senshuu meeru-o atarashii ofisu-de isha-ga yukkuri kaita youda.
last week email-Acc new office-in doctor-Nom slowly wrote it appears

All the target sentences had the following structure: *time expression, NP1, adjective, location, NP2, modifier (noun or adjective), verb, auxiliary verb*. The same 8 reversible and 8 non-reversible verbs in the previous tasks were used to make the target sentences. Each verb was used twice to create 32 canonical sentences in total. The pairs of the sentences that shared the same verb were split into two item lists, and the scrambled sentences which were made from the canonical sentences on one list were added to the other list, and vice versa, for the purpose of counterbalancing. Forty-eight filler sentences with various structures were added to each item list. Thus, each item list consisted of 8 each of the reversible canonical/scrambled sentences, 8 each of the non-reversible canonical/scrambled sentences, and 48 filler sentences, for a total of 80 sentences.

All the words used in the sentence items, except for common Japanese names, were chosen from *Nakama 1 & 2 2nd Edition* (Hatasa, Hatasa, & Makino, 2009). Common Japanese names were used for about half of the animate nouns for the purpose of minimizing the repetition of the same nouns in the task. One of the nouns in each reversible sentence was a regular noun (such as “girl” and “older brother”) while the other noun was a name. In the effort to minimize the effect of this regular noun/name contrast on the participants’ performance, regular nouns were nominative-marked in half of the reversible sentences and names were nominative-marked in the other half. Half of the non-reversible sentences had regular nouns as their subjects, and the other half had names as subjects. The sentence items were presented in Japanese orthography with *furigana* (*hiragana* as phonetic guide) over *kanji*.

3.7.3 Procedure

The participants were assigned to one of the two counter-balanced item lists. Prior to the task, the participants were given a word list and were instructed to study the words in the list. The word list included all of the vocabulary items in the target sentences and about half of the vocabulary items in the filler sentences. Glosses in English were provided next to each Japanese word. The purpose of the word list was to alleviate the effect of unfamiliar vocabulary items on the participants’ sentence comprehension.

The sentence items were presented with the DMDX software (Forster & Forster, 2003). The sentences were presented using the non-cumulative self-paced moving-window technique (Just, Carpenter, & Woolley, 1982). The display was initially filled with dashes replacing the characters of the entire sentence to be presented. When the participant pressed the designated key, the first word of the sentence appeared replacing the dashes corresponding to the word.

When the key was pressed again, the second word appeared and the previous word was replaced with dashes. The participant read the sentences this way until the end, by pressing the key at their own pace. Ten practice items preceded the actual trials.

In order to minimize situations in which a participant might blindly press the key without reading the sentences for meaning, all of the sentence items, including the filler items, were followed by a comprehension question about the sentence. The comprehension questions consisted of a question statement and two choices of answers, and the participant was asked to choose the correct answer by pressing the corresponding key. The questions on the target sentences were about either the first or second NP, and only one of the nouns in the two choices of answers actually appeared in the preceding sentence. For the purpose of encouraging the participants to read the sentences for meaning, instead of just scanning the words to answer the comprehension questions, the questions on the filler sentences consisted of various types, and two words that actually appeared in the preceding sentence were often used for the two answer choices.

3.7.4 Analysis

The data analysis included the reading times (RT) of the second word (NP1, i.e., nominative or accusative NP), the fifth word (NP2, i.e., accusative or nominative NP), and the seventh word (Verb). In order to take the spillover effect into consideration, the RT data of second & third words (NP1 + 1), fifth & sixth words (NP2 + 1), and seventh & eighth words (Verb + 1) were also analyzed. The analyses of the “NP1 + 1”, “NP2 + 1”, and the “Verb + 1” positions will be provided in the subsequent results section only if they were substantially different from those of the “NP1”, “NP2”, and “Verb” positions. Data from all the participants were included in the analysis. The participant who made errors most frequently with the comprehension questions of the target items still showed 75% accuracy.

The data analysis was based on the residual RT (F. Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). In the present analysis, the residual RT was calculated first by computing a linear regression equation as a function of the word length and the raw RT by each participant. The predicted RT for each data point was calculated by substituting the word length into the equation, and the residual RT was calculated by subtracting the predicted RT from the raw RT. Due to the mixing of the different scripts in the Japanese writing system, the reading time of a word was expected to be influenced by both the number of morae and characters. Therefore, following Mazuka, Itoh, & Kondo (2002), the number of morae as well as the number of characters were used to define the word length. For instance, for the word "新聞を" ('newspaper-Acc'), there are 5 morae (shi-n-bu-n-o) and 3 characters, and therefore, the word length was defined as 4 (the average of the two values). All of the sentences, including the fillers, were included for the calculation of the linear regression equation for each participant. However, the first word (time expression) and the eighth word (auxiliary verb) were excluded for the calculation because the same vocabulary items were repeated several times at these word positions throughout the task.

3.8 Results of Task 3

The residual RT data were analyzed using linear mixed-effects modeling in R. The data from the different participant groups were analyzed separately. The fixed effects were Reversibility and Word order, and the random effects were Subjects and Items. A mixed-effects model with the interaction of Reversibility and Word order was applied to the datasets first. If the analysis did not indicate a significant interaction of the two fixed effects, the linear mixed-effect model without the interaction term was applied to obtain the results. For the purpose of observing the effect of Word order for the reversible and non-reversible sentences, however, the results based on the model with the interaction will also be reported even when there was no significant interaction. The results using a random slopes model will be reported if the model was justified by the data and if the results were substantially different from those without random slopes.

Trials in which an error was made were omitted from the analysis of RT. Also, some participants occasionally pressed a wrong key as they read the sentences. In such occasions, the entire sentence was removed from the analysis. Residuals that were 2.5 standard deviations away from the linear mixed-effects model were trimmed. The residual RT at each position by participant group is provided in Table 3.1. The number of data points trimmed (using the linear mixed-effects model without random slopes) is shown in Table 3.2.

Table 3.1: Mean residual reading times for each sentence type by participant group in Task 3

		NP1		NP1+1		NP2		NP2+1		Verb		Verb+1	
		Can	Scr	Can	Scr	Can	Scr	Can	Scr	Can	Scr	Can	Scr
ADV	Rev	227	301	94	378	211	252	381	445	-331	-227	-1001	-799
	NRev	431	-8	340	-95	-99	424	-207	260	-306	-268	-965	-876
HI	Rev	439	355	259	237	70	140	441	461	-362	-359	-1287	-1207
	NRev	345	28	428	-78	-183	284	-240	109	-400	-343	-1347	-1180
LI	Rev	309	433	2	269	336	320	637	758	-484	-626	-1745	-1866
	NRev	580	-71	294	-340	-171	451	-306	461	-602	-464	-1793	-1674
NS	Rev	13	1	31	-13	-22	28	-83	-19	-97	-90	-216	-196
	NRev	22	10	91	43	-46	8	-121	-40	-81	-81	-183	-189

Table 3.2: The number of trimmed data points for each analysis in Task 3

	Total	NP1	NP1 + 1	NP2	NP2 + 1	Verb	Verb + 1
ADV	622	19 (3.05)	13 (2.09)	11 (1.77)	18 (2.89)	25 (4.02)	16 (2.57)
HI	416	11 (2.64)	9 (2.16)	12 (2.88)	12 (2.88)	13 (3.13)	15 (3.61)
LI	590	14 (2.37)	13 (2.2)	13 (2.2)	15 (2.54)	12 (2.03)	9 (1.53)
NS	615	12 (1.95)	17 (2.76)	13 (2.11)	19 (3.09)	13 (2.11)	17 (2.76)

Note: The numbers in the Total column are the total number of items that were considered for analysis. The numbers in parentheses show the percentages of the items trimmed.

3.8.1 The ADV group

The mean residual RT of Task 3 by the ADV group is presented in Figure 3.3.

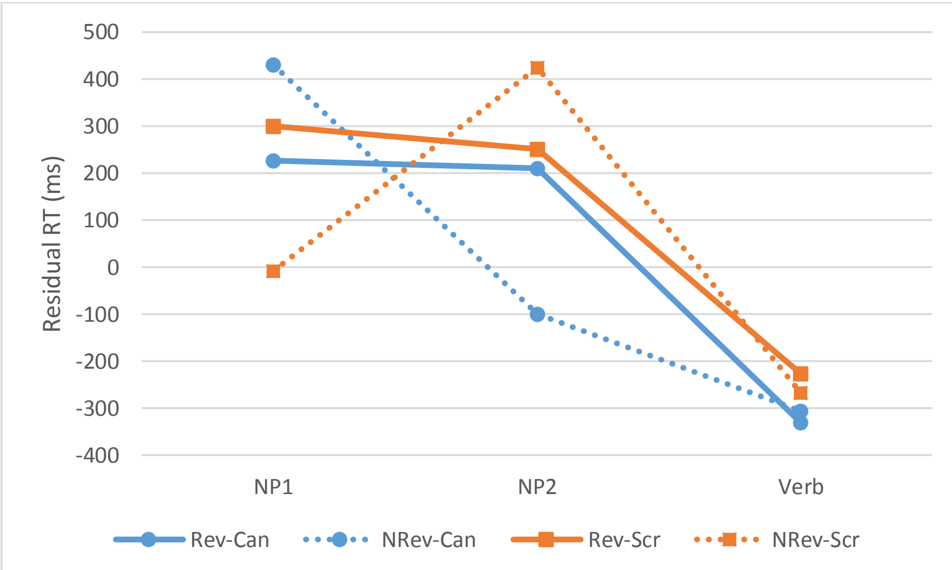


Figure 3.3: Mean residual reading times for the ADV group in Task 3

For the NP1 position, the interaction of Reversibility and Word order was significant (t = 5.042, p < 0.0001). The main effect of Word order was significant (t = 3.647, p = 0.0003), but the main effect of Reversibility was not significant (t = 0.651, p = 0.52). For the reversible sentences, the effect of Word order was not significant (t = 0.979, p = 0.3281), but it was significant for the non-reversible sentences (t = 6.193, p < 0.0001). On the other hand, the analysis of the residual RT at the "NP1 + 1" position indicated that the effect of Word order was significant for the reversible sentences (t = 2.641, p = 0.0085) where the mean residual RT for the scrambled sentences were significantly longer than that of the canonical sentences.

For the NP2 position, the interaction of the two fixed effects was significant (t = 4.541, p < 0.0001). The main effect of Word order was significant (t = 5.203, p < 0.0001) while the main effect of Reversibility was not significant (t = 0.899, p = 0.376). For the reversible sentences, the effect of Word order was not significant (t = 0.497, p = 0.6193) whereas it was significant for the non-reversible sentences (t = 6.955, p < 0.0001).

For the Verb position, the main effect of Word order was significant (t = 2.392, p = 0.0171) while the main effect of Reversibility was not significant (t = 0.196, p = 0.8459). The interaction was not significant ($\chi^2(1)$ = 1.2559, p = 0.2624). An analysis using a model with the interaction term indicated that the effect of Word order was significant for the reversible sentences (t = 2.471, p = 0.0138) while it was not significant for the non-reversible sentences (t = 0.934, p = 0.351).

To summarize the results for the ADV group, a substantial difference was observed between the reversible and non-reversible sentences. For the reversible sentences, the effect of Word order was observed at the Verb and the "NP1 + 1" positions. For the non-reversible sentences, on the other hand, the residual RT of the scrambled condition was significantly shorter at NP1 but the result was reversed at NP2.

3.8.2 The HI group

The mean residual RT by the HI group is shown in Figure 3.4. For the NP1 position, the main effects of Reversibility as well as Word order were significant (Reversibility: t = 2.253, p = 0.0346; Word order: t = 2.283, p = 0.0231). The interaction was not significant ($\chi^2(1)$ = 1.8199, p = 0.1773). An analysis using a model with the interaction term indicated that the effect of Word order was significant for the non-reversible sentences (t = 2.596, p = 0.0099), but it was not significant for the reversible sentences (t = 0.63, p = 0.5288). On the other hand, with an analysis of the same dataset using a random slopes model (by-subject for the effect of Reversibility and by-item for the effect of Word order), all of the significant differences above turned non-significant.

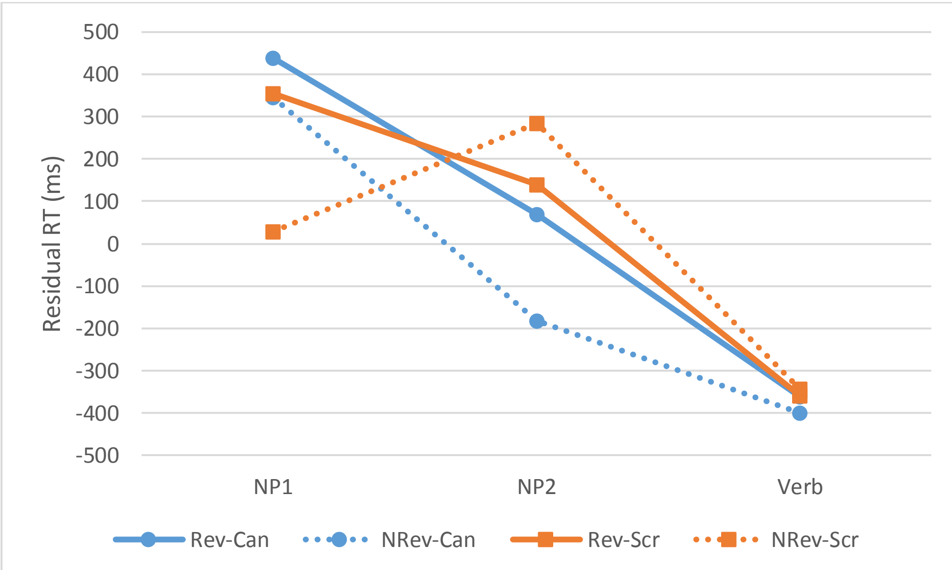


Figure 3.4: Mean residual reading times for the HI group in Task 3

The analysis of the residual RT at the "NP1 + 1" position indicated a significant interaction of the two fixed effects (t = 2.143, p = 0.0328), which turned non-significant using a random slopes model. The effect of Word order for the non-reversible sentences remained significant.

For the NP2 position, the interaction of the fixed effects was significant (t = 2.744, p = 0.0064). The main effect of Word order was significant (t = 3.545, p = 0.0004) while the main effect of Reversibility was not significant (t = 0.527, p = 0.6019). The effect of Word order was not significant for the reversible sentences (t = 0.675, p = 0.5001) but it was significant for the non-reversible sentences (t = 4.563, p < 0.0001). On the other hand, with an analysis using a random slopes model (by-subject and by-item for the effect of Word order), the interaction turned non-significant (t = 1.456, p = 0.1556) while the effect of Word order for the non-reversible sentences remained significant (t = 2.279, p = 0.0293).

For the Verb position, neither of the main effects was significant (Reversibility: $t = 0.08$, $p = 0.937$; Word order: $t = 0.507$, $p = 0.613$). The interaction was not significant ($\chi^2(1) = 0.3687$, $p = 0.5437$). At the "Verb + 1" position, on the other hand, the main effect of Word order was significant ($t = 2.282$, $p = 0.0231$). An analysis using a model with the interaction term indicated that the effect of Word order was significant for the non-reversible sentences ($t = 2.074$, $p = 0.0388$), but it was not significant for the reversible sentences ($t = 1.136$, $p = 0.257$).

To summarize, the mean residual RT by the HI group indicated a pattern quite similar to that by the ADV group. There was little effect of Word order for the reversible sentences. On the other hand, for the non-reversible sentences, the mean RT of the scrambled sentences was significantly shorter at the NP1 position and the result was reversed at the NP2 position.

3.8.3 The LI group

The mean residual RT by the LI group is shown in Figure 3.5.

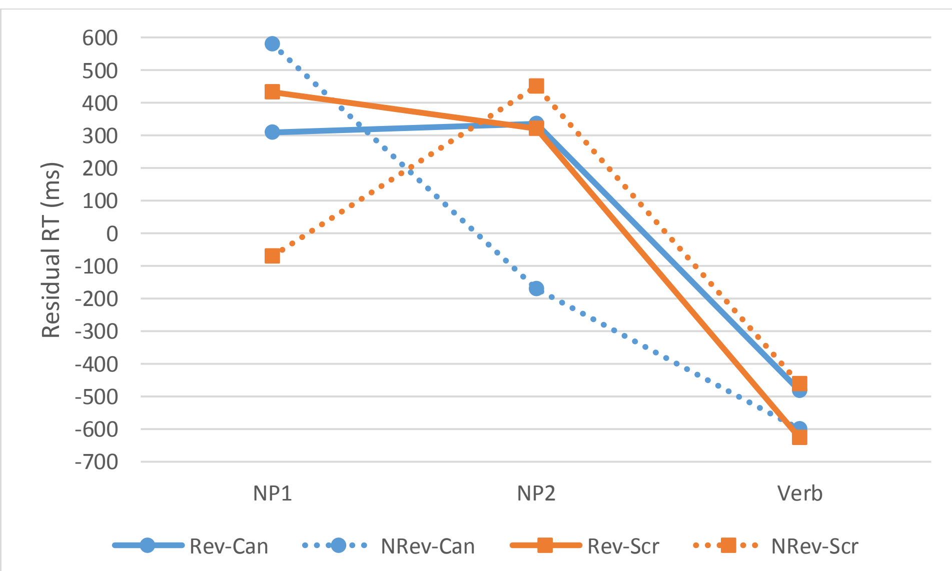


Figure 3.5: Mean residual reading times for the LI group in Task 3

For the RT of the NP1 position, the interaction of the two fixed effects was significant (t = 4.391, p < 0.0001). The main effect of Word order was significant (t = 3.111, p = 0.002), but the main effect of Reversibility was not significant (t = 0.859, p = 0.3971). The effect of Word order was significant for the non-reversible sentences (t = 5.337, p < 0.0001) while it was not significant for the reversible sentences (t = 0.914, p = 0.3613).

For the NP2 position, the interaction was significant (t = 3.673, p = 0.0003). The main effect of Word order was significant (t = 3.389, p = 0.0008), but the main effect of Reversibility was not significant (t = 1.014, p = 0.3186). The effect of Word order was significant for the non-reversible sentences (t = 5.02, p < 0.0001) while it was not significant for the reversible sentences (t = 0.203, p = 0.8389).

For the Verb position, the interaction of Reversibility and Word order was significant (t = 3.356, p = 0.0009). The main effects of Reversibility and Word order were not significant (Reversibility: t = 0.327, p = 0.746; Word order: t = 0.201, p = 0.841). The effect of Word order was significant for the reversible sentences (t = 2.131, p = 0.0336) as well as for the non-reversible sentences (t = 2.618, p = 0.0091) although the effect of Word order was opposite for the reversible and non-reversible sentences. At the “Verb + 1” position, the effect of Word order turned non-significant for both sentence types (Reversible: t = 1.208, p = 0.2274; Non-reversible: t = 1.889, p = 0.0594).

The results above indicate a similar pattern to those of the ADV and HI groups. For the NP1 and NP2 positions, no significant effect of Word order was observed for the reversible sentences. On the other hand, there was a significant effect of Word order for the non-reversible sentences. At the NP1 position, the mean residual RT was significantly shorter for the scrambled sentences, and the result was opposite for the NP2 position.

3.8.4 The NS group

The mean residual RT by the NS group is shown in Figure 3.6.

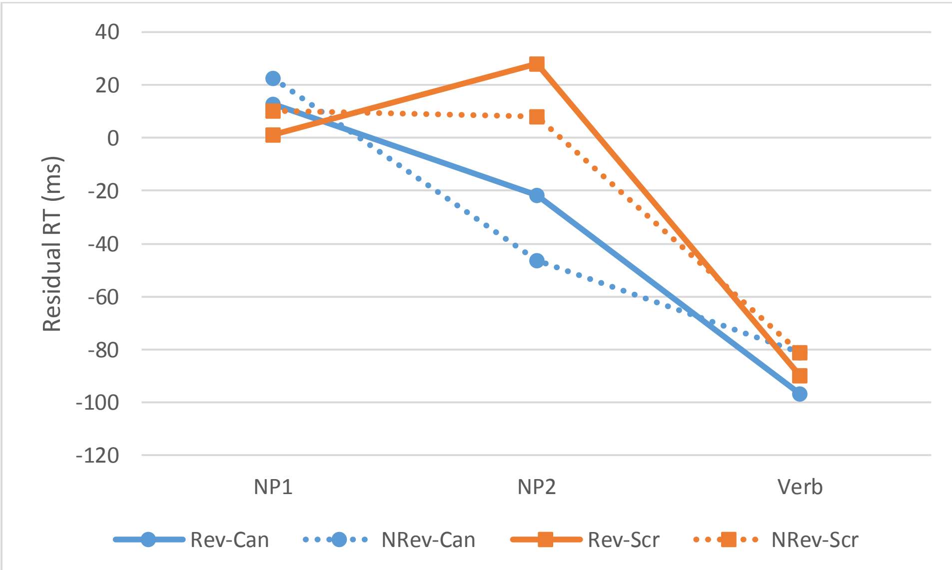


Figure 3.6: Mean residual reading times for the NS group in Task 3

For the RT of the NP1 position, none of the main effects for Reversibility or for Word order were significant (Reversibility: t = 0.608, p = 0.544; Word order: t = 0.883, p = 0.378). The interaction was not significant ($\chi^2(1) = 0$, p = 0.9986).

For the NP2 position, the main effect of Word order was significant (t = 3.284, p = 0.0011) while the main effect of Reversibility was not significant (t = 1.127, p = 0.2686). The interaction was not significant ($\chi^2(1) = 0.0055$, p = 0.9407).

For the Verb position, the main effect of Reversibility as well as that of Word order were not significant (Reversibility: t = 1.108, p = 0.2765; Word order: t = 0.387, p = 0.6992). The interaction was not significant ($\chi^2(1) = 0.3745$, p = 0.5405).

The results above replicated the finding of the previous studies with native speakers of Japanese (e.g., Mazuka et al., 2002) which observed a significant effect of word order at the second NP position but not at the first NP. No significant effect of Reversibility was observed.

3.8.5 Discussion

As indicated above, the analysis of the NS data replicated Mazuka et al.'s (2002) results. The slowdowns in reading the scrambled sentences were observed at the second NP, suggesting that this was where the reanalysis of the sentence structure took place. That no significant difference was observed at the verb position seems to support this view. The NS data also indicated that the slowdowns at the second NP position were significant for both reversible and non-reversible sentences. It is thus speculated that the slowdowns were due to syntactic reasons rather than lexical/semantic reasons.

The results from the three NNS groups were strikingly similar. For the reversible sentences, no significant effect of Word order was observed either at the NP1 or NP2 position (except for the "NP1 + 1" position for the ADV group). For the non-reversible sentences, on the other hand, the first NPs of the scrambled sentences were read significantly faster than those of the canonical sentences, and the result was reversed at the NP2 position.

At first sight, the slowdowns at the NP2 position in the non-reversible scrambled sentences seem to resemble the NS data. However, because the first NPs of the scrambled sentences (i.e., accusative-marked NP) were read significantly faster than the first NPs of the canonical sentences (i.e., nominative-marked NP), the slowdowns at the NP2 position in the non-reversible sentences are most likely to be due to the lexical items in the sentences. That is, accusative-marked inanimate NPs were read faster than nominative-marked animate NPs. For the purpose of avoiding the repetition of the words, the animate nouns in the Task 3 stimuli included Japanese names as well as some nouns that might not have been as familiar to the NNS participants (e.g., *lawyer*, *Chinese person*). Thus, it is possible that the animate nouns in the stimuli were generally more difficult for the NNS participants to recognize than the inanimate nouns.

Taken together, the results indicated that the processing of the scrambled sentences by the three NNS groups were similar, and the NNS participants did not seem to engage in native-like syntactic processing when they read the scrambled sentences.

3.9 General discussion

The result of the pilot study using the sentence correctness decision task indicated that the scrambled sentences caused longer reaction times and lower accuracy rates among the NS and NNS participants, which suggests that the NNS participants experienced additional psychological cost in reading and comprehending scrambled sentences as did the NSs. On the other hand, the results of Task 3 (self-paced reading task) indicated that the processing of scrambled sentences by L2 learners might be quite different from that of NSs. While the NS participants experienced slowdowns at the NP2 position when they read the OSV sentences, the NNS data did not provide evidence for such slowdowns. Thus, the results of Task 3 are consistent with the SSH's claim that L2 sentence processing may not involve a native-like syntactic processing.

If we assume that L2 learners engage in native-like syntactic processing at a high proficiency level as observed in previous studies (e.g., Hopp, 2006; Steinhauer et al., 2006), there would be at least two reasons why Task 3 failed to observe evidence for such native-like processing. The first reason is the way the sentence stimuli were set up. Mazuka et al. (2002) did not observe a significant difference in RT between their simple canonical and scrambled sentences in which the two NPs were adjacent to each other. When there was a modifier phrase between the two NPs, however, they observed a significant scrambling effect. Thus, for the purpose of observing a clearer scrambling effect, two-word location phrases were inserted between the two NPs in

Task 3. While the results of Task 3 indicated a scrambling effect among the NS participants, insertion of the location phrases might not have been optimal in observing a scrambling effect among the NNS participants. As observed, Hara (2009) found that the L1 Korean learners of Japanese slowed down at the gap position when they read the "short scrambling" ditransitive sentences in which the scrambled accusative NP was adjacent to the dative NP. On the other hand, the same Korean participants did not experience slowdowns when they read the "long scrambling" sentences in which there was a three-word location phrase inserted between the two NPs whereas the NS participants slowed down in reading such sentences. The gap processing becomes more costly and thus takes longer when the gap is further away from its antecedent (Gibson, 1998). As L2 sentence processing is expected to be more demanding than its L1 counterpart in terms of cognitive resources (e.g., working memory) it requires, the NNS participants in Task 3 might not have been able to engage in syntactic processing of the scrambled sentences due to the relative complexity of the sentence stimuli.

The second reason for the potential failure to observe the scrambling effect in Task 3 may be the participants. In those studies which observed native-like sentence processing by L2 learners, the participants were highly proficient in their target language. Thus, while the ADV participants in the present study scored higher in the proficiency test than the others, their proficiency on average might not have been high enough to engage in native-like sentence processing. To explore this possibility, the RT data of the 10 participants who scored the highest in the SPOT test (scores ranging from 111 to 121 out of 125) were analyzed. However, the result was similar to that of the ADV group, and slowdowns were not observed at the NP2 position of the reversible scrambled sentences. (The residual RT of this analysis is presented in Appendix 3B.)

One interesting point to note in the data from the ADV group is that, significant slowdowns were observed at the "NP1 + 1" and the Verb positions when they read the reversible scrambled sentences. As noted, because both nominative and accusative NPs are animate in reversible sentences, the only information signaling that the sentence is scrambled are the case markers. Since the HI and LI data did not indicate significant slowdowns at these positions in the reversible scrambled sentences, the slowdowns among the ADV participants may indicate that they integrated the case-marker information more consistently than the HI and LI participants. The significantly longer RT at the Verb position may suggest that the filler-gap integration took place at this position when the ADV participants read the reversible scrambled sentences. It is also worth noting that the HI and LI data indicated significant slowdowns at the verb in the non-reversible scrambled sentences. Due to the "inanimate-animate" NP order, it might have been easier for the participants to identify scrambling when they read the non-reversible scrambled sentences. Thus the slowdowns at the Verb position among the HI and LI participants may be taken as evidence that the filler-gap integration took place at this position. On the other hand, the data from the two groups did not indicate slowdowns at the verb when the participants read the reversible scrambled sentences (for the LI group, the reversible scrambled sentences were read faster at the verb than the canonical sentences). This might suggest that, contrary to the ADV participants, the HI and LI participants might not have identified scrambling when they read the reversible scrambled sentences.[35]

While the results of Task 3 indicated that the processing of scrambled sentences by L2 learners might be different from that of NSs, the results are not definitive at all given that Hara

[35] The comprehension questions in Task 3 did not require the identification of nominative/accusative NP (i.e., which of the two NPs in a given sentence was subject or object) to choose the correct answer. In order to observe whether L2 learners can reliably use case-marker information to identify the grammatical case of NPs in reversible sentences, future investigation should include comprehension questions that require such an identification.

(2009) observed a native-like slowdown pattern among advanced L1 Korean learners. Investigation with different types of stimuli (possibly ones with a shorter filler-gap distance) and with different participant groups in terms of proficiency and L1 are necessary to better understand how scrambled sentences are read and comprehended by L2 learners. Studies that examine brain responses (ERP, fMRI) also appear quite promising. For instance, Kim et al. (2009), using fMRI, observed more activity in some areas of the brain (i.e., the left inferior frontal gyrus and the left dorsal prefrontal cortex) during NS comprehension of simple transitive scrambled sentences than during their comprehension of canonical sentences. Such a method, coupled with behavioral investigations, would provide further insights into the processing of scrambled sentences by L2 learners.[36]

[36] While the present analysis did not examine each participant's performance on Task 3 in relation to the tasks in Chapter 2, such an analysis may provide important information about how individuals' grammatical knowledge, production performance, and comprehension of OSV sentences are related. A future analysis will address this issue.

CHAPTER 4: SYNTACTIC PERSISTENCE IN THE PRODUCTION OF SCRAMBLED SENTENCES BY L2 LEARNERS

4.1 Introduction

The results of the picture description task (Task 2) indicated that some of the NNS participants experienced difficulties in producing scrambled sentences. While such difficulties seemed to stem mostly from a lack of grammatical knowledge, the results also indicated that grammatical knowledge does not necessarily ensure the correct production of scrambled sentences, especially the reversible ones. The results of the reading tasks in Chapter 3 further suggested that L2 learners, especially those with lower proficiency, might not integrate case markers effectively in identifying scrambling in OSV sentences. Thus, effective practice methods to facilitate processing of scrambled sentence are called for. As processing (i.e., production and comprehension) of scrambled sentences requires that learners be conscious of the function of the postpositional particles, such processing practice may also contribute to the acquisition of the particles.

The tasks in the present chapter were aimed at exploring whether the production of scrambled sentences might be facilitated by using the phenomenon of syntactic persistence. Syntactic persistence refers to a tendency for speakers to reuse a previously produced or comprehended sentence structure for the processing of subsequent sentences. The phenomenon has been employed extensively in investigations of the cognitive mechanism of language processing, and it has been suggested that syntactic persistence may reflect language learning processes (e.g., Chang, Dell, & Bock, 2006). Studies that investigated L2 sentence production using syntactic persistence have yielded promising results, showing the positive contribution of syntactic persistence in the acquisition of L2 grammatical structures (e.g., McDonough & Kim, 2009; McDonough & Mackey, 2008). Hence, if the scrambling structure persists among L2 learners of Japanese, it may open the possibility of using the method in scaffolding the learning of scrambling.

This chapter will first provide a general review of experimental studies that investigated the phenomenon of syntactic persistence. It will then review L2 studies (between L1 and L2, and within L2) that employed the syntactic persistence technique followed by studies on Japanese scrambling. The presentation of the present study will follow.

4.2 Literature review

4.2.1 Syntactic persistence phenomenon

Syntactic persistence (syntactic priming) refers to the phenomenon in which language users employ syntactic structures to which they have recently been exposed in producing subsequent sentences, when there is an alternative structure to express the same idea. For instance, someone who has just heard the sentence "The student gave an apple to the teacher." is more likely to say "The lady bought an orange for the boy." than "The lady bought the boy an orange." when asked to describe a picture in which a lady is paying for an orange and giving it to a boy. In this case, the prepositional-object dative structure (V NP PP; henceforth PO) just heard has *persisted* into the subsequent sentence production. The picture has been described using the PO structure even though the same idea could have been expressed with double-object dative (V NP NP; henceforth DO).

Bock (1986) was the first study to investigate syntactic persistence employing a systematic experimental design. The structures Bock examined were datives (PO/DO) and transitives (active/passive). Under the guise of a memory task, the researcher asked the participants to repeat sentences out loud (the prime sentences) and then describe pictures (the target sentences). As expected, there was a significantly more frequent production of POs after the participants had heard and repeated PO primes and more frequent production of DO datives after exposure to DO primes. Active and passive primes also significantly increased the production of active and passive targets, respectively.

In subsequent studies, Bock and colleagues further examined the triggers of syntactic persistence. Bock (1989) demonstrated that PO primes with *to* or *for* equally facilitated the production of the *to*-dative, suggesting that syntactic persisitence was not just the repetition of closed-class function words such as prepositions. Bock and Loebell (1990) observed that syntactic persistence is not sensitive to thematic structures (*The wealthy widow gave her Mercedes to the church* primed *The wealthy widow drove an old Mercedes to the church*). However, it is sensitive to phrase structure (*Susan brought a book to Stella* did not prime *Susan brought a book to study*). This finding also seems to suggest that syntactic persistence is not sensitive to overall sound or prosody of the sentence. Further, Pickering, Branigan, and McLean (2002) observed that shifted POs (e.g., *The patient showed to the doctor his leg*) did not prime PO datives, and therefore, syntactic persistence appears to be sensitive to constituent order.

While initial syntactic persistence studies looked at L1 English, subsequent studies observed syntactic persistence in Dutch (Hartsuiker & Westenberg, 2000) and German (Scheepers, 2003) as well as within L2 and between L1 and L2. Syntactic persistence has also been observed in structures other than datives and transitives: in noun phrases (Cleland & Pickering, 2003), complementizer *that* (V. Ferreira, 2003), and order of auxiliaries (Hartsuiker & Westenberg, 2000). Additionally, syntactic persistence appears to take place from production to production (e.g., Bock, 1986) as well as from comprehension to production (e.g., Branigan, Pickering, & Cleland, 2000).

One interesting finding in syntactic persistence is *lexical boost* (Pickering & Branigan, 1998), that is, when the prime and the target have the same head lexical item (verb or noun), the priming is enhanced. Based on findings on lexical boost combined with the observation that priming is not affected by variations in tense, aspect, or number morphemes, Pickering and Branigan (1998) proposed an integration of syntactic information into a detailed network model by Roelofs (1992, 1993) and Levelt, Roelofs, & Mayer (1999). Roelofs and Levelt et al. assume three strata in lexical production: the conceptual stratum (in which concepts are generated), the lemma stratum (in which the base form of words such as *give* and *send* is represented), and the word-form stratum (in which the morpho-phonological information of words is specified).

Pickering and Branigan (1998) proposed that lemma nodes in the lemma stratum are connected to nodes that represent the syntactic properties of the words (*combinational nodes*). For instance, ditransitive verbs such as *give* and *send* are linked to the *NP, NP* node (representing DO datives) as well as to the *NP, PP* node (representing PO datives). Based on this model, when someone produces (or hears) a sentence such as "The boy gave an apple to the teacher.", the *give* lemma node and the *NP, PP* combinational node are temporarily activated. In a subsequent production of a dative sentence, the already activated *NP, PP* node (rather than the *NP, NP* node) is more likely to be used. Thus, according to this model, syntactic persistence takes place due to residual activation of the combinational node by the prime.

The model correctly predicts lexical boost. When the prime includes a certain verb and a certain syntactic structure, the corresponding lemma and combinational nodes as well as the links between them are activated. Due to such activation, a higher magnitude of priming is expected when the target includes the same verb. However, syntactic priming can still be observed when a prime and a target have different verbs. Because combinational nodes are shared between lemma nodes, use of a certain syntactic structure activates the corresponding combinational node, which increases the likelihood of the same structure being used in the target. Finally, priming is not sensitive to the morphology of the verb because the lemmas are directly linked to the combinational nodes.

While the residual activation account of Pickering and Branigan's model matches the empirical data obtained from research studies, another important feature of syntactic persistence, which has been confirmed in several experimental studies, may not be readily explained with such an account. That is, syntactic persistence is very long-lasting (Bock, Dell, Chang, & Onishi, 2007; Bock & Griffin, 2000; Boyland & Anderson, 1998; Branigan, Pickering, Stewart, & McLean, 2000). For instance, Bock and Griffin (2000) tested the syntactic persistence of datives and transitives with varying numbers of items intervening between primes and targets, and the results indicated that persistence did not decay even with ten items intervening. Boyland and Anderson (1998) observed long-lasting persistence of syntactic structures (datives and transitives) over 20 minutes. Observations such as these have led some researchers to speculate that syntactic persistence reflects a rather long-term implicit learning mechanism which "may undergo fine-tuning in every episode of adult language production" (Bock & Griffin, 2000: 189). According to the implicit learning account, processing a specific structure strengthens the mapping between the features of messages to be expressed and the particular syntactic structure. This makes the structure more accessible when messages with similar features are to be expressed in the future.

To reconcile the empirical evidence that supports the implicit learning account as well as the residual activation account, V. Ferreira and Bock (2006) suggested that multiple mechanisms contribute to syntactic persistence. That is, syntactic persistence reflects long-term, implicit structural learning mechanisms, but it can also be facilitated by the activation of specific lexical items. Hartsuiker, Bernolet, Schoonbaert, Speybroeck, and Vanderelst (2008) observed results that support Ferreira and Bock's multi-factorial account, showing that lexical boost is short-lived while syntactic persistence itself lasts longer, regardless of the modality of prime and target (written or spoken). Based on the results, Hartsuiker et al. maintain that the implicit learning mechanisms are the main source of syntactic persistence, and the lexically-driven mechanism contributes to the greater magnitude of syntactic persistence.

Although exactly which mechanism is responsible for syntactic persistence is still under debate, researchers generally seem to agree that, for syntactic persistence within L1 at least, what actually persists is something like phrase structure rules (Branigan, 2007; Pickering & V. Ferreira, 2008). First, a phrase structure rule such as VP → V NP PP is independent of specific lexical items. Therefore, syntactic persistence takes place even when different verbs or prepositions are used (Bock, 1989). The insensitivity of syntactic persistence to thematic structures (Bock & Loebell, 1990) is consistent with phrase structure analysis as well because phrase structure rules do not encode thematic roles. Additionally, syntactic persistence seems to be insensitive to empty categories or argument/adjunct distinctions, which are not specified in phrase structure rules (e.g., Bock & Loebell, 1990). Phonological closeness may also be irrelevant to syntactic persistence. Cleland and Pickering (2003) found that phonologically close words (e.g., ship vs. sheep) did not enhance the priming effect. Syntactic persistence, on the other hand, is sensitive to word order (Pickering et al., 2002) just as phrase structure rules specify word order (e.g., VP → V NP PP but not VP → V PP NP).

However, the priming of thematic structures seems to require further investigation (Branigan, 2007; Pickering & V. Ferreira, 2008). While Bock and Loebell (1990) found that priming is insensitive to thematic structures, Chang, Bock, and Goldberg (2003) found evidence that the order of thematic roles may contribute to syntactic persistence. Chang et al. manipulated the primes so that the order of the thematic role is switched without changing the syntactic structure of the sentence (*The maid rubbed polish onto the table* [theme-location] vs. *The maid rubbed the table with polish* [location-theme]). It was found that, for instance, *The maid rubbed polish onto the table* primed *The farmer heaped straw onto the wagon*, but not *The farmer heaped the wagon with straw*, suggesting the influence of thematic structure on syntactic persistence.

4.2.2 L2 studies of syntactic persistence

Broadly speaking, there appear to be two major directions in syntactic persistence research in L2: 1) investigation into the mechanism of L2 production/comprehension, and 2) acquisition of new syntactic structures in L2. The former is mostly studied through syntactic persistence between L1 and L2, and seems to be more concerned with investigating the model of L2 production/comprehension in relation to the L1, that is, whether syntax is shared between L1 and L2, and if so what is shared and what is not. The latter is largely studied through syntactic persistence within L2 and seems to be more focused on the function of syntactic persistence as a tool to measure and facilitate L2 learning.

4.2.2.1 Cross-linguistic studies

As mentioned above, one strand of L2 syntactic persistence research is concerned with the representation of L2 syntax in relation to its L1 counterpart. The primary question is whether syntax is shared between L1 and L2 (shared-syntax account) or separate (separate-syntax account). One obvious advantage of shared syntax is reduced redundancy, because, according to the account, bilinguals store rules that are the same in two languages only once. Sharing L1 and L2 syntax may also be advantageous for those bilinguals who code-switch between two languages because they would not have to access two separate syntactic representations as they frequently switch languages in a conversation (Hartsuiker, Pickering, & Veltcamp, 2004). On the other hand, the separate-syntax account is motivated by the fact that, although the rules to form a particular syntactic structure may be largely similar in any two given languages, there are often minute but critical differences in the formation. (Hartsuiker et al. employ English and Spanish passives to make the point.) Therefore, having two syntactic representations between L1 and L2 may facilitate more efficient and accurate processing. In addition, it may be more efficient for bilinguals who use only one language at a time (Hartsuiker et al., 2004).

The empirical evidence obtained from cross-linguistic syntactic persistence studies seems to favor the shared syntax hypothesis. For example, Loebell and Bock (2003) examined priming of datives and transitives among L1 German-L2 English speakers. The results showed that there was a significant priming of datives between the two languages (in both directions; PO priming was weaker, however), but priming was not observed for transitives in either direction. On the other hand, Hartsuiker et al. (2004) observed priming of transitives in Spanish-English bilinguals. The result showed that the participants produced significantly more passives following passive primes than following active or intransitive sentences, indicating a significant priming effect.

Meijer and Fox Tree (2003) found priming of datives among Spanish-English bilinguals. They employed the sentence recall method in which the participants first read an English sentence, then read a Spanish prime, and finally are asked to remember the first English sentence. The researchers found that the participants falsely remembered English DO sentences as PO datives more frequently after Spanish PO dative primes than after non-PO primes. Salamoura and Williams (2007) also found priming of datives from Greek (L1) to English (L2) in a sentence completion task. Additionally, Schoonbaert, Hartsuiker, and Pickering (2007) observed priming of datives within L1 (Dutch) and L2 (English) as well as between L1 and L2 (both directions). Cross-linguistic syntactic persistence has been observed in structures other than transitives and datives. Bernolet, Hartsuiker, and Pickering (2007) found priming of noun phrases from German to Dutch (but not between Dutch and English), and Desmet and Declercq (2006) observed priming of relative-clause attachment (high vs. low) from L1 Dutch to L2 English.

Just as there is a lexical boost within L1, there appears to be a translation-equivalence boost (the prime and the target have translation-equivalents). As mentioned earlier, Schoonbaert et al. (2007) examined the priming of datives within Dutch (L1) and English (L2) as well as between the two languages. While they observed a lexical boost within L1 and L2 as expected from the previous studies, they found a translation-equivalence boost from L1 to L2 but not from L2 to L1. Schoonbaert et al. speculate that the asymmetry may be due to the different strengths of links between lexical representations and concepts in L1 and L2. As the Revised Hierarchical Model by Kroll and Stewart (1994) states, the link between the L2 lexical representation and the concept may be less strong than its L1 counterpart, and this weaker link may be responsible for the absence of a translation-equivalence boost from L2 to L1.

Studies have also found that cross-linguistic syntactic persistence is highly sensitive to word order. Bernolet et al. (2007) observed priming of NPs (*a blue baby* vs. *a baby that is blue*) within Dutch and English among L1 Dutch learners of English. However, they did not find priming of NPs between the two languages. According to Bernolet et al., the word order in relative clauses differs between Dutch and English – the word order equivalent of *a baby that is blue* in Dutch is *a baby that blue is* – and the difference in word order probably prevented priming from taking place. Indeed, in an experiment which tested NP priming between Dutch and German, which have an identical relative clause word order, they observed clear priming effects. As mentioned earlier, Loebell and Bock (2003) observed priming of datives (especially DO) between German (L1) and English (L2) but failed to find such an effect with transitives. Loebell and Bock attribute the absence of priming to different passive word orders between German and English. While the *by*-phrase is placed at the end of a passive clause in English, its German equivalent precedes the past participle. The sensitivity of cross-linguistic syntactic persistence to word order seems to echo the findings in L1 studies (e.g., Pickering et al., 2002).

Salamoura and Williams (2007) replicated Pickering et al. (2002) with L1 Greek learners of English. They found that, while PO and DO L1 primes facilitated the production of PO and DO structures respectively in L2, shifted PO in L1 (e.g., *the boy gave to the teacher an apple*) did not prime PO structure in L2. This suggests that it is indeed the constituent order that persists across languages. In a subsequent experiment in the same study, Salamoura and Williams further tested the priming of four different structures: PO, DO, locative, and “provide with” structures. The results indicated that the “provide with” structure primed DOs and that locatives primed POs.

Because the first post-verbal arguments of "provide with" structures and DOs, as well as those of locatives and POs, coincide in syntactic structure and thematic role (NP-recipient/goal for the former and NP-theme for the latter), the researchers speculate that the shared syntactic structure and thematic role in the first post-verbal argument may trigger priming between the prime and the target.

With respect to the issue of word order and syntactic persistence as well as the asymmetry in syntactic persistence between L1 and L2, Park (2007) provides informative findings. In the investigation of syntactic persistence among L1 English-L2 Korean and L1 Korean-L2 English participants, Park found that 1) there was a significant priming effect for transitive structures (active/passive) except for priming from L2 Korean to L1 English, and 2) there was little priming effect for ditransitive structures (DO/PO) except the priming from L1 to L2 in the L2 English group. Thus, it appears that syntactic persistence was more robust in the "L1 to L2" direction than the "L2 to L1" direction in this study. The results also seem to suggest that, between English and Korean at least, exactly the same surface word order is not required for syntactic persistence to take place. English is a head-initial language whereas Korean is a head-final language. If exactly the same word order was a requirement for syntactic persistence, the priming in transitive structures would not have been observed in the study. Syntactic persistence from L1 Korean to L2 English was also observed in Shin (2008).

Based on findings in cross-linguistic syntactic persistence studies, Hartsuiker et al. (2004) proposed a bilingual lexical-syntax model, which derives from the monolingual model by Pickering and Branigan (1998). In this model, the lemmas for an L1 verb and its L2 equivalent are connected to the same combinational nodes, and the combinational nodes are shared by different verbs. A prime in one language (a PO sentence, for instance) will activate the lemma node as well as the *NP, PP* combinational node. Therefore, when asked to describe a dative picture in the other language, the PO structure is more likely to be used, because of the residual activation of the *NP, PP* node. A translation-equivalence boost is correctly predicted in this model because the translation-equivalent words share a conceptual node. Schoonbaert et al. (2007) modified Hartsuiker et al.'s model and reduced the strength of the link between L2 lemmas and conceptual nodes, based on the absence of a translation-equivalence boost from L2 to L1 in their study.

To summarize, cross-linguistic priming studies have observed syntactic persistence between languages, which suggests that syntactic representations may be shared between languages. However, in order for syntactic persistence to take place, it appears that the order of constituents needs to be shared between prime and target, while the structure of the entire clause may not need to be identical (Salamoura & Williams, 2007). Although research so far has concentrated on priming between typologically similar languages, Park (2007) and Shin (2008) observed priming between English and Korean, suggesting that syntactic representations can be shared between typologically different languages as well.

4.2.2.2 Within-L2 studies

Another strand of L2 syntactic persistence research has examined priming within L2, and its focus has been the investigation of the function of syntactic persistence as a tool to measure and facilitate L2 learning. Syntactic persistence studies in L1 are almost always concerned with alternatives that are equally acceptable in the language – for example, active/passive, DO/PO dative. For L2 learners, on the other hand, the alternation can be between developmentally advanced and simple structures or between accurate target structures and inaccurate structures

(McDonough & Kim, 2009). If syntactic persistence reflects an implicit learning mechanism, as Bock and Griffin (2000) suggest, it may be used as a device to guide L2 learners from a simple or inaccurate L2 syntax stage to a more advanced/accurate L2 syntax stage.

McDonough (2006) investigated whether syntactic persistence of English dative constructions is observed among L2 learners of English, the majority of whom were L1 Chinese speakers. The data were obtained using the confederate scripting technique, in which two participants describe pictures to each other while one of them is actually a confederate who reads primes aloud from a scripted list (Branigan et al., 2000). The results indicated that syntactic priming occurred for the PO datives while the priming of DO datives was not observed, regardless of whether participants repeated the primes orally. The author suggested that the reason for the absence may be that L2 learners' use of DO datives might be specific to certain verbs. In other words, they might not have acquired the abstract syntactic information for the DO structure that could have been activated by primes. However, the study did observe some contribution of the priming technique to the production of DO structures. A post-hoc analysis indicated that when the participants were exposed only to DO primes (Experiment 2), the proportion of DO datives produced with specific verbs (*ask* and *teach*) decreased. In addition, among the eight verbs that had not been used to produce DO datives in the previous experiment, six were used at least once. The results seem to indicate that greater type frequency (i.e., opportunities to apply different lexical items to form a specific structure) facilitates the use of DO datives in L2 English learners.

L2 learners' dependence on lexical items in forming certain syntactic structures and the importance of type frequency in acquiring more advanced structures has been confirmed in subsequent studies. Kim and McDonough (2008) observed that, in all three levels of EFL learners (L1 Korean), learners produced significantly more passives when the prime and the target had the same verb. This tendency, however, was stronger for the low-proficiency learners, suggesting that their passive formation might be more dependent on specific verbs than that of more advanced learners. McDonough and Kim (2009) investigated whether type frequency in primes and the learners' prompts (i.e., words in the materials to elicit production) influence L1 Thai EFL learners' production of *wh*-questions. The results indicated that participants who received high type-frequency prompts produced correct *wh*-questions with higher frequency than those who received low type-frequency prompts. This confirmed the importance of lexical diversity for promoting learners' production of certain grammatical structures. Unexpectedly, however, prime type-frequency did not seem to affect learners' *wh*-question production.

While McDonough and Kim (2009) did not directly investigate syntactic priming of *wh*-questions and, therefore, it is not clear whether priming takes place in EFL learners' *wh*-question formation, McDonough and Mackey (2008) clearly observed such priming effects. In the latter study, the partipants were L1 Thai college students, and more advanced English learners at the same university served as their interlocutors. The results showed that participants produced more advanced *wh*-questions after hearing their interlocutors' utterance of advanced *wh*-questions. The results also showed that participants who produced more advanced *wh*-questions after their interlocutors' primes produced more advanced *wh*-questions in the post-test, which took place three and seven weeks later. The post-hoc analysis indicated that, among those who primed more frequently to the interlocutors' primes, the ones who moved to a higher question-formation stage produced *wh*-questions with more lexical diversity.

Shin (2008) evaluated the effectiveness of syntactic persistence in the learning of datives and phrasal verbs among L1 Korean learners of English, using a pre- and post-test design (Experiment 3). The results indicated that the participants' performance on grammaticality judgment and picture description improved after the syntactic persistence training, suggesting that the technique was effective for the learning of the L2 structures. Shin also tested whether explicit knowledge instruction (i.e., grammar instruction) combined with syntactic persistence was more effective than syntactic persistence alone. The results indicated that the explicit learning condition can be more beneficial in the short-term learning of the L2 structures although implicit learning can be as beneficial in the long run.

To summarize, studies have found within-L2 syntactic persistence in different structures: datives (McDonough, 2006; Shin, 2008), passives (Kim & McDonough, 2008), phrasal verbs (Shin, 2008), and *wh*-questions (McDonough & Mackey, 2008). In addition, Bernolet et al. (2007) observed priming of noun phrases (*a blue baby* vs. *a baby that is blue*) among L1 Dutch learners of English. While the aforementioned studies were conducted in L2 English, Flett (2006) observed the priming of VS order among L1 English learners of Spanish. Thus, syntactic persistence is certainly not limited to English as an L2. Similarly to their L1 counterparts, within-L2 studies have observed lexical boost (Kim & McDonough, 2008) and have observed that syntactic persistence within L2 is also long-lasting (McDonough & Mackey, 2008). On the other hand, within-L2 syntactic persistence seems to be different from the L1 counterpart in that the structures that are still underdeveloped in the learners' interlanguage system are less likely to persist (McDonough, 2006).[37] Perhaps the most important contribution of within-L2 studies is the finding on the facilitative effects of syntactic persistence on the development of L2 syntax. Exposure to more advanced structures as primes (with an increased type frequency) appears to scaffold the learning of such structures (McDonough & Kim, 2009; McDonough & Mackey,

[37] Morishita, Satoi, and Yokokawa (2010) also report smaller priming effects in the DO structure among lower proficiency L1 Japanese learners of English. Similarly to McDonough (2006), they attribute the smaller priming effects to the lack of abstract syntactic representation.

2008; Shin, 2008). This suggests that syntactic persistence may be used as a tool to promote L2 learners' acquisition of L2 syntactic structures.

<u>4.2.3 Syntactic persistence studies on Japanese scrambling</u>

So far as the author is aware, there have been only two studies that investigated syntactic persistence on Japanese scrambling, and both studies observed priming of word order among NSs of Japanese.

Yamashita, Chang, and Hirose (2002) examined the syntactic persistence of dative sentences, using a sentence repetition task (Potter & Lombardi, 1998). The participants read sentences on a computer monitor, and were asked to produce them after a distractor task. The target sentences had the *wa-o-ni* order as shown in (4.1). The target sentences were preceded by three types of prime sentences as shown in (4.2): a) *wa-o-ni* order, b) *wa-ni-o* order, and c) *wa-ni-o* order but the –*ni* marked NP is an adjunct rather than an argument (*wa*-adjunct_*ni*-*o*).

(4.1) *Target*:

Taro-wa	hana-o	onnanoko-ni	okutta.
-Top	flower-Acc	girl-Dat	sent

'Taro sent flowers to the girl.'

(4.2) a. *wa-o-ni prime*:

Akiko-wa	kagi-o	tomodati-ni	ageta.
-Top	-Acc	friend-Dat	gave

'Akiko gave the key to the friend.'

b. *wa-ni-o prime*:

Akiko-wa	tomodati-ni	kagi-o	ageta.
-Top	friend-Dat	key-Acc	gave

c. *wa-adjunct_ni-o prime*:

Akiko-wa	sinya-ni	kagi-o	ageta.
-Top	midnight-at	key-Acc	gave

'Akiko gave the key at midnight.'

The results indicated that the participants produced *wa-ni-o* sentences more frequently after the *wa-ni-o* primes than after the *wa-o-ni* primes. It was also found that *wa-ni-o* sentences were produced more frequently after the *wa*-adjunct_*ni-o* primes than after the *wa-o-ni* primes, suggesting that syntactic persistence is sensitive to the surface order of arguments in Japanese as well (cf. Bock & Loebell, 1990).

Tanaka (2008) investigated syntactic persistence of the OSV word order, using priming from comprehension to production (Experiment 3). The participants first saw a prime picture on a computer monitor, and were asked to judge whether a prime sentence subsequently presented described the prime picture correctly. The participants then saw a target picture and were asked to describe the picture. All of the target pictures had inanimate agent and animate patients (e.g., "the screwdriver poking the conductor"). The results indicated that, although the overall production rate of OSV sentences was low, the participants produced more OSV sentences after the OSV primes (10.5%) than after the SOV primes (6%).

4.3 The present study

The present study attempted to examine whether the Japanese OSV word order would persist among L2 learners. While the main task (Task 4) used regular SOV/OSV sentences as primes, the smaller-scale follow-up study (Task 6) used questions in SOV/OSV orders as primes. In addition, Task 6 attempted to examine whether postpositional particles might prime, using results of Task 5 (a fill-in-the-blank task) as baselines.

As reviewed, it has been suggested that syntactic persistence reflects long-term, implicit structural learning mechanisms (Bock & Griffin, 2000). It has also been suggested that syntactic persistence can facilitate the learning of abstract L2 syntactic structures (McDonough, 2006;

Shin, 2008). Thus, if the scrambling structure persists among L2 learners of Japanese, it will suggest the possibility that the method can be used as an instructional tool to facilitate the learning of scrambling word order. Task 6 also examined whether a syntactic persistence would be beneficial in the learning of postpositional particles.

As with other tasks in the study, the participants' performance was observed by proficiency groups. This was to examine whether general proficiency in the target language might interact with syntactic persistence, since previous studies (e.g., McDonough, 2006; Morishita et al., 2010) have found smaller priming effects among learners with lower proficiency. The two reversible sentence types (reversible and non-reversible) were also included for the purpose of observing how the additional difficulty experienced with the reversible scrambled sentences in the picture description task (Task 2) might interact with syntactic persistence.

4.4 Method for Task 4

4.4.1 Participants

The 54 NNS and 20 NS participants in other tasks in the present study participated in Task 4.

4.4.2 Materials and design

There were four types of prime sentences. They were all simple transitive sentences as shown below. (See Appendix 4A for a list of prime sentences in Task 4.)

(4.3) a. *Reversible canonical*:
Mika-ga Jiro-o karakatta.
-Nom -Acc teased.
'Mika teased Jiro.'

b. *Reversible scrambled*:
Jiro-o Mika-ga karakatta.
-Acc -Nom teased

c. *Non-reversible canonical*:
Ryohei-ga pasokon-o tsukatta.
-Nom computer-Acc used
'Ryohei used the computer.'

d. *Non-reversible scrambled*:
Pasokon-o Ryohei-ga tsukatta.
-Acc -Nom used

In order to make the prime sentences, 8 verbs that can be used in reversible sentences and 8 verbs that can be used in non-reversible sentences were chosen. Each verb was used twice to make 32 canonical sentences in total. The pairs of the sentences that shared the same verb were split into two item lists, and the scrambled sentences which were made from the canonical sentences on one list were added to the other list, and vice versa, for the purpose of counterbalancing. Forty filler sentences of various structures were added to each item list. The sentences were digitally recorded by a native speaker of Japanese.

All the words used in the sentence items, except for common Japanese names, were chosen from *Nakama 1 & 2 2nd Edition* (Hatasa, Hatasa, & Makino, 2009). Common Japanese names were used for about half of the animate nouns for the purpose of minimizing the repetition of the same nouns in the task. Half of the reversible sentences had regular nouns (e.g., "elderly man" and "waiter") for both nominative and accusative NPs, and the other half had names for both nominative and accusative NPs. Half of the non-reversible sentences had regular nouns as their subjects, and the other half had names as subjects.

The set of pictures used in Task 2 was also used in Task 4 (16 reversible and 16 non-reversible). As with Task 2, the persons and the inanimate objects in the pictures were labeled with words so that the participants could use those words for picture description. The labels for the agent and patient were placed at the same height in each picture. Also, the label for the agent was placed on the left side for half of the pictures and on the right side for the other half. The label for the verb was placed at the bottom of each picture. Figure 4.1 shows the examples of the reversible and non-reversible pictures.

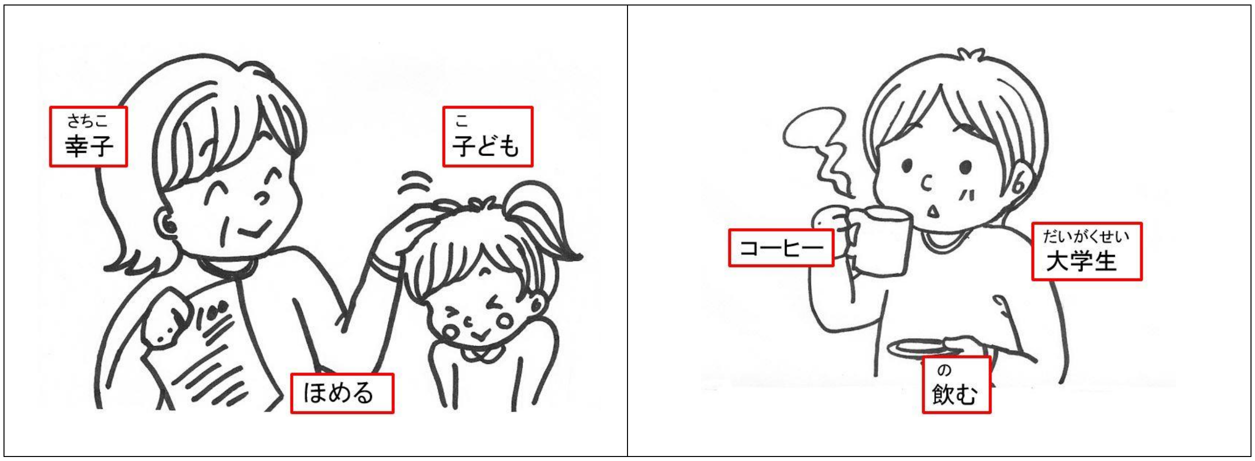


Figure 4.1: An example of a reversible picture (left) and a non-reversible picture (right) used in Task 4

The 16 reversible pictures were coupled with the 8 reversible canonical and 8 reversible scrambled prime sentences within each item list. Likewise, the 16 non-reversible pictures were coupled with the non-reversible canonical and scrambled prime sentences. Forty pairs of sentences and pictures were also included in each list as fillers. Therefore, each item list consisted of 8 each of reversible canonical/scrambled items and non-reversible canonical/scrambled items as well as 40 filler items, totaling 72 items. The recordings of the sentences and the pictures were inserted into Microsoft PowerPoint slides. The items in each list were pseudo-randomized so that the items of the same condition would not be adjacent to each other. In an effort to alleviate the influence of the item presentation order, two pseudo-randomized versions were created for each item list.

4.4.3 Procedures

The procedures were modeled after Bock (1986). The participants were first given a word list and were instructed to study the words in the list. The word list included all of the vocabulary items that were in the target items (prime sentences and pictures) as well as about half of the vocabulary items that were in the filler items. Glosses in English were provided next to each Japanese word. In addition to checking the meaning of the words, the NNS participants were also asked to pronounce each word aloud, while the NS participants were only asked to study the words in the list. The purpose of the word list was to alleviate the effect of unfamiliar words or unfamiliar pronunciation of the words on the participants' performance during the syntactic persistence task.

Next the participants received a folder which included 38 sentences and 38 pictures, and they were asked to study the items in the folder for about 10 minutes so that they could recognize the items during the task. The sentences and pictures in the folder were those of the filler items.

The participants were then told that the following task would be a recognition task, and that they would be asked to indicate whether the sentences and the pictures they would encounter during the task were among those in the folder that they had just studied. The purpose of the memory task was to minimize the participants' attention to their speech production. They were also told that they would be asked to listen to and repeat the sentences aloud and describe the pictures.

In order to alleviate the effect of the previous picture description task (Task 2) on the present task, the participants were assigned to an alternative list of items. That is, if a participant was prompted to describe a picture in the OSV order in Task 2, the participant was assigned to the other item list in which the same picture was preceded by an SOV prime sentence, and vice versa, in Task 4. Each participant worked on the task individually in a sound-resistant computer booth. The presentation of the material was carried out on a laptop computer with Microsoft PowerPoint. Each trial began with the presentation of a sound icon and a sentence. The prime sentences were written out, as several pilot trials indicated that L2 learners experienced difficulties repeating the sentences from aural input alone. The participants were instructed to listen to the sentence by clicking the sound icon and to repeat the sentence aloud. If they felt they had not repeated the sentence fluently, they were instructed to listen to the sentence again and repeat it. They were asked to try this as many times as necessary until they were able to repeat the sentence fluently. After the repetition of the sentence, they were asked to indicate whether the sentence was in the folder by saying “yes” or “no” aloud. In the next slide, a picture was presented, and the participants were asked to describe the picture in a full sentence, using all the words in the picture. After the picture description, they were again asked to say “yes” or “no” aloud to indicate whether the picture was in the folder. A visual example of a trial is shown in Figure 4.2.

There were four practice items prior to the actual trial. The experimenter went through the practice items with the participant to make sure that he/she could perform the task as intended. The participants’ responses were recorded with a digital recorder and were transcribed for scoring.

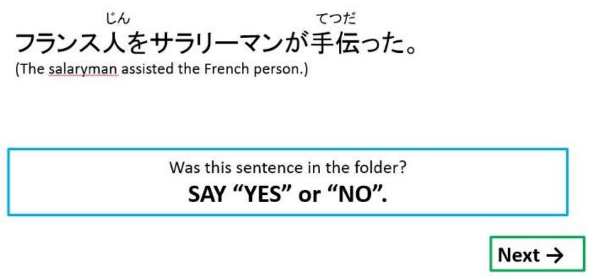
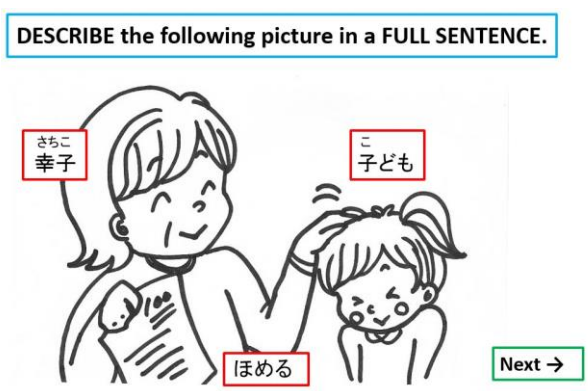
Visual presentations	Participants’ tasks
	1. The participants listened to the prime by clicking the sound icon. (The prime was written out along with an English translation as shown.) 2. The participants repeated the prime sentence aloud.
	3. The participants indicated whether the sentence was in the folder they had studied by saying “yes” or “no”.
	4. The participants described the picture in a full sentence using all the words in the picture.
	5. The participants indicated whether the picture was in the folder they had studied by saying “yes” or “no”.

Figure 4.2: Visual presentation of a trial in Task 4

4.4.4 Scoring and analysis

The transcribed responses were coded either “Canonical”, “Scrambled”, “Other”, or “NA”. A response was coded as “Canonical” when it correctly described the picture in the SOV word order. A response was considered “Scrambled” when it correctly described the picture in the OSV word order. The particles *–ga* and *–wa* were considered subject markers and the particle *–o* was considered as an object marker in this analysis. As with Task 2, however, there were quite a few instances in which the object was marked with the *–ni* particle, especially in the descriptions of the “reversible” pictures. Therefore, *–ni* was also considered as an object marker although the marking of the object NP with *–ni* made the responses ungrammatical. Since the topic marker –*wa* can replace the object marker, the “NP1-*wa* NP2-*ga* V” responses were also classified as Scrambled.

Responses were classified as “Other” in the following instances: a) when the descriptions contradicted the pictures (e.g., *Daigakusei-o coffee-ga nonda.* ‘The coffee drank the college student.’), b) when wrong particles were used making the descriptions ungrammatical (e.g., double subject or double object sentences), or c) when unintended structures were used (e.g., passives). The “NA” coding was given to responses when the participants did not repeat the preceding prime sentences correctly. In most cases, repetition errors were caused by omission or replacement of particles. There were 48 instances (out of 1728 opportunities) of repetition errors for the NNS participants, and one instance of repetition error (out of 640) for the NS participants. The proportions of the Canonical, Scrambled, and Other responses for each participant group and for each condition, after the removal of the “NA” items, are provided in Appendix 4B.

For the statistical analysis, the responses that were coded as Scrambled were given the score of 1, and the responses coded as Canonical and Other were given the score of 0. The utterance rates of scrambled sentences in each condition were analyzed by each participant group using linear mixed-effects modeling in R. The fixed effects were Reversibility and Word order, and the random effects were Subjects and Items. A mixed-effects model with the interaction of Reversibility and Word order was applied to the datasets first. If the analysis did not indicate a significant interaction of the two fixed effects, the linear mixed-effects model without the interaction term was applied to obtain the results. The results using a random slopes model are reported if the model was justified by the data and if the results were substantially different from those without random slopes.

4.5 Results of Task 4

The utterance rates of scrambled sentences by condition and by participant group are provided in Figure 4.3.

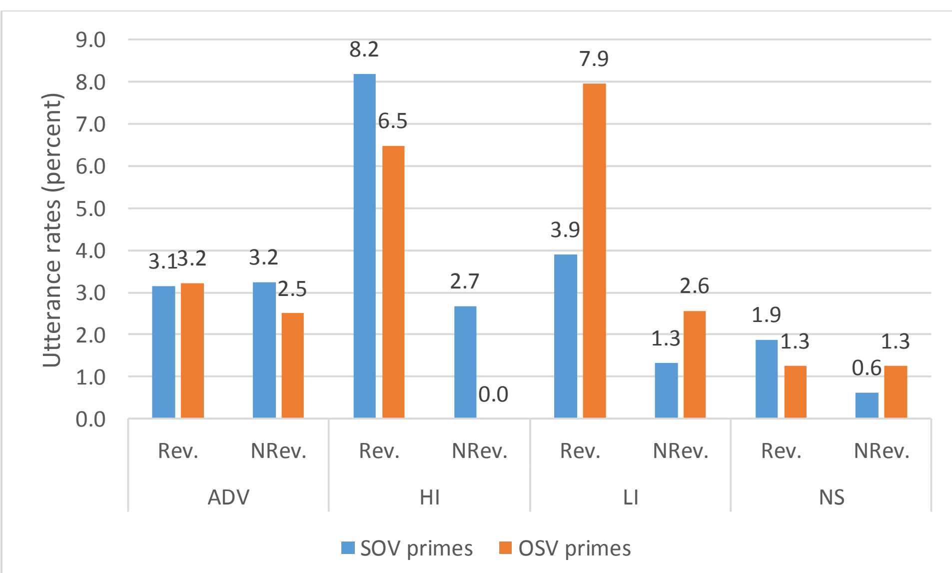


Figure 4.3: Utterance rates of scrambled sentences in Task 4

For the ADV group, the main effects of Reversibility and Word order were not significant (Reversibility: $t = 0.21$, $p = 0.835$; Word order: $t = 0.353$, $p = 0.724$). The interaction of the two fixed effects was not significant ($\chi^2(1) = 0.1011$, $p = 0.7505$).

For the HI group, the main effect of Reversibility was significant ($t = 3.17$, $p = 0.0035$) while the main effect of Word order was not significant ($t = 1.19$, $p = 0.2347$). The effect of Word order was opposite of the expected direction (i.e., more scrambled sentences were produced after the SOV primes). The interaction of the two fixed effects was not significant ($\chi^2(1) = 0.0963$, $p = 0.7563$). When the same dataset was analyzed using a model with random slopes (by-subject for the effect of Reversibility), the main effect of Reversibility was not significant ($t = 1.567$, $p = 0.1361$).

For the LI group, the main effect of Reversibility was significant ($t = 2.938$, $p = 0.0034$) while the main effect of Word order did not reach significance ($t = 1.829$, $p = 0.0679$). The interaction of the two factors was not significant ($\chi^2(1) = 0.9079$, $p = 0.3407$). When the dataset was analyzed using a model with the interaction term, the effect of Word order was marginally significant for the reversible sentences ($t = 1.964$, $p = 0.05$) while it was not significant for the non-reversible sentences ($t = 0.625$, $p = 0.532$).

For the NS group, the main effects of Reversibility and Word order were not significant (Reversibility: $t = 0.688$, $p = 0.497$; Word order: $t = 0.001$, $p = 0.999$). The interaction of the two fixed effects was not significant ($\chi^2(1) = 0.5244$, $p = 0.469$).

4.5.1 Discussion

The results of Task 4 overall did not show syntactic persistence of the OSV word order. The only exception is the reversible condition among the LI participants, which observed a marginally significant effect of syntactic persistence. In addition to the general absence of syntactic persistence, the overall production rates of scrambled sentences in Task 4 were quite low. Thus, it has to be concluded that a syntactic persistence task in a format as in Task 4 may not be an effective instructional tool to elicit production of OSV sentences.

The general absence of syntactic persistence of the OSV structure as well as the low overall OSV production rates may be attributed to the infrequent occurrence of OSV sentences. According to the implicit-learning account of syntactic persistence, processing a specific structure strengthens the mapping between the features of messages to be expressed and the particular syntactic structure, which makes the structure more accessible when messages with similar features are to be expressed in the future. As reviewed in Chapter 1, the occurrence of OSV sentences is rather rare, especially in written contexts, and for this reason, it is possible that the mapping between the message (the transitive structure with an agent and a patient, in this case) and the abstract representation of the OSV structure is very weak among L2 learners as well as NSs. By the same token, it is possible that the mapping between the "transitive" message structure and the canonical SOV structure was so prominent among the participants that a subtle "push" such as syntactic persistence might not have been strong enough to cause changes in their choice of the syntactic structures.

Tanaka (2008) observed a significant syntactic persistence of the OSV word order among NSs of Japanese while Task 4 did not. In addition, the production rate of OSV sentences among the NS participants in Task 4 was much lower than that of Tanaka's. One factor that is speculated to have caused the difference, aside from the way the primes were presented, is the target pictures. In Tanaka, all of the target pictures had inanimate agents and animate patients (e.g., "the screwdriver poking the conductor") while the agent was always animate in the target pictures in

Task 4. Given the prominence of animate entities over inanimate ones, the target pictures in Tanaka's study might have motivated the participants to place the animate patients at the beginning of their sentence production. In Task 4, the agent was always animate in the target pictures and was often drawn larger than the patient. Thus, the participants in Task 4 might have been less motivated to produce OSV sentences when they described the target pictures.

A few other factors might have contributed to the lack of syntactic persistence and the infrequent OSV production in Task 4. One potential factor is the way the prime sentences were presented. Since it was found, during the pilot trial phase, that L2 learners had difficulty repeating the prime sentences from aural input alone, the written-out prime sentences were presented as the participants listened to the sentences. This might have allowed the participants to simply read the written-out sentences aloud, making the processing of the prime sentences "shallower" than it needed to be to have the syntactic persistence effect. Another potential factor is the lack of motivation for producing scrambled sentences (other than the OSV primes) in Task 4. As observed in Chapter 1, production of scrambled sentences can be motivated by such factors as "referentiality" to the preceding context and "heaviness" of the NPs (Yamashita, 2002). Since such motivational factors to produce scrambled sentences were not included in Task 4, the participants might have simply used the canonical SOV order when they were asked to describe the transitive pictures.

While the previous studies (e.g., McDonough, 2006; Morishita et al., 2010) have found smaller priming effects among learners with lower proficiency, Task 4 observed a significant OSV priming effect only in the reversible condition among the LI participants. This contradictory result in Task 4 might be explained if we assumed, using Pickering and Branigan's (1998) model, that the connection between the lemma nodes of the reversible verbs and the combinational node of the "Subject-Object" word order is less established among the LI participants. As observed in Task 2 (cf. Figure 2.11), the accuracy rates among the LI participants during their production of reversible sentences were rather low. Thus, it is speculated that their knowledge of how to formulate reversible sentences was underdeveloped, which might have given "room" for the OSV priming sentences to take effect. The higher overall production rate of reversible OSV sentences among the HI group might be explained with the same reason. On the other hand, for the NS and ADV individuals, the connection between the reversible verbs and the canonical "Subject-Object" combinational node might have been too solid for the syntactic persistence of the OSV structure to take place.

4.6 The follow-up tasks (Tasks 5 and 6)

The results of Task 4 indicated that the task was not effective in eliciting OSV sentences. To further explore the possibility of using syntactic persistence as a means to elicit the correct production of OSV sentences among L2 learners, Task 6 employed questions in OSV word order as primes (e.g., "Whom who saw?") to elicit scrambled sentences. Japanese is a *wh-in-situ* language and allows multiple *wh*-elements in a clause. In Task 6, the participants listened to a question either in SOV or OSV order, and were asked to answer the question based on the information on the picture that was simultaneously presented to them.

There are several reasons to use the "question as primes" format. First, it has been commonly discussed that the scrambled element at the beginning of a clause receives some emphasis or focus (e.g., Shibatani, 1990; Yamashita, 2002). Therefore, asking a question in OSV order, and thus placing some emphasis on the accusative *wh*-NP might motivate participants to describe the pictures with the accusative NP first. The phrase order in questions can also indicate what the questioner is really interested in, which may facilitate the responses starting with the object. Second, syntactic persistence studies using the confederate scripting technique (e.g., Branigan et al., 2000; McDonough, 2006) have found robust priming effects, indicating that syntactic persistence takes place in dialogue. It has also been suggested that syntactic persistence in dialogue reflects alignment between interlocutors (Pickering & Garrod, 2004). Thus, it was expected that a syntactic persistence task in a "Q & A" format might elicit more robust priming effect. Third, during the debriefing interview after the second experimental session, some participants said that they felt awkward when they described the pictures (e.g., "The college student drinks coffee.") since they would seldom say such a sentence in Japanese, by itself, in a natural setting. A few participants also expressed that they felt OSV sentences were less formal than SOV sentences, and hence, and felt somewhat "inappropriate" to say such sentences in an experimental setting. Therefore, it was hoped that the "Q & A" format would bring some "naturalness" to the task by providing some context to participants' responses. Finally, since the prime questions shared the same verbs as the target pictures, a more robust priming effect was expected through *lexical boost* (Pickering & Branigan, 1998).

While the primary purpose of Task 6 was to investigate whether the "questions as primes" format might elicit production of OSV sentences, another purpose of the task was to observe whether the questions would facilitate the correct choice of postpositional particles among L2 learners. As observed in Task 1, many participants in the HI and LI groups seemed to have experienced difficulty choosing the correct *–o* particle to mark the accusative NPs in the reversible sentences. Because the questions in Task 6 included particles which the participants could use to provide correct responses, whether such questions could improve L2 learners' use of particles was of interest. If improvement in L2 learners' use of particles is observed in the task, the question format could be easily adopted in classroom instruction. In addition to the canonical reversible sentences, six sentences with different particles were included to examine whether the improvement in the use of particles, if any, might be extended beyond the –*o* particle.

The NNS participants worked on Task 5, a fill-in-the-blank task similar to Task 1, before Task 6. The purposes of Task 5 were to evaluate the participants' knowledge of scrambling at the time of their participation, and to assess the participants' knowledge of particles, which were to be compared with their use of particles in Task 6.

4.7 Method for the follow-up tasks

4.7.1 Participants

The participants in this follow-up portion of the study were 30 L2 learners and 16 native speakers of Japanese. The NNS individuals participated in both Tasks 5 and 6 while the NSs only participated in Task 6. Due to the limited availability of potential participants, those who had already participated (or were going to participate) in the previous tasks were invited to participate in this portion of the study. About half of the participants in these two tasks had gone through the debriefing interview for the previous tasks, and the other half worked on Tasks 5 and 6 as the third study session before the debriefing interview. In order to minimize the effect of the debriefing interview for those who had already gone through the procedure, the participants were invited to Tasks 5 and 6 only if it had been more than two months since they had been debriefed.

For the purpose of grouping, the NNS participants who had already participated in the debriefing interview were asked to take Versions B and A of the SPOT test again, as it had been more than three months since most of the participants took SPOT at the beginning of the first study session. Those NNS participants who were invited to Tasks 5 and 6 for the third session did not take SPOT again, and their original SPOT scores were used for grouping.

Based on the sum of the SPOT version B and A scores, the NNS participants were grouped into either the Advanced (Adv) or Intermediate (Int) group. There were 16 participants in the Adv group and 14 participants in the Int group. Out of the possible 125 points, the SPOT scores of the Adv participants ranged from 98 to 122 (mean: 109.5), and the scores of the Int participants ranged from 45 to 89 (mean: 66.3).

4.7.2 Materials and design

4.7.2.1 Task 5

The materials in Task 5 were identical to those in Task 1 except that the following 6 items were included in place of the locative scrambled items in Task 1. The 6 items correspond to filler pictures in Task 2 with which the participants appeared to have experienced difficulties in terms of the correct choice of particles. These items were included in Task 5 to evaluate the participants' knowledge of the particles and to provide a baseline for Task 6. (The particles in the parentheses indicate the common errors that the participants made in describing the pictures in Task 2).

(4.4) a. Hanako-ga omocha-de (*-o, -to) asonda.
-Nom toy-with played
'Hanako played with the toy.'

b. Tsuyoshi-ga isu-ni (*-de, -o) suwatta.
-Nom chair-on sat
'Tsuyoshi sat on the chair.'

c. Kenji-ga hoteru-ni (*-de, -o) tomatta.[38]
 -Nom hotel-at stayed
 'Kenji stayed at the hotel.'

d. Mai-ga basu-ni (*-o) notta.
 -Nom bus-on rode
 'Mai rode the bus.'

e. Ichiro-ga kaisha-de (*-ni, -o) hataraita.
 -Nom company-at worked
 'Ichiro worked at the company.'

f. Ayako-ga roku-kiro-Ø (*-o, -ni) yaseta.
 -Nom 6-kilogram lost
 'Ayako lost 6 kilograms.'

Each of the two counterbalanced lists included 8 each of reversible canonical/scrambled sentences and non-reversible canonical/scrambled sentences, 6 sentences in (4.4) above, and 42 filler sentences, for a total of 80 sentences. As with Task 1, the item lists were presented in a paper-and-pencil format. In an effort to alleviate the influence of the item presentation order, two pseudo-randomized versions were created for each item list.

4.7.2.2 Task 6

The picture items in Task 6 were identical to those in Task 4. Six pictures that corresponded to the sentences in (4.4) were treated as target items in Task 6 while they were used as filler items in Task 4. The six pictures are shown in Appendix 4C.

While the primes in Task 4 were simple transitive sentences in canonical and scrambled word orders, the primes in Task 6 were questions about the pictures in canonical and scrambled word orders. The examples of primes for each condition are shown in (4.5).

[38] Communication with several native speakers of Japanese indicated that some of them felt that the *–de* particle is also acceptable in this sentence. However, the item was included in the analysis because none of the NS participants in Task 6 used the *–de* particle to produce the sentence.

(4.5) a. *Reversible canonical:*
Dare-ga dare-o karakaimashita-ka?
who-Nom who-Acc teased-Q
'Who teased whom?'

b. *Reversible scrambled*:
Dare-o dare-ga karakaimashita-ka?
who-Acc who-Nom teased-Q

c. *Non-reversible canonical:*
Dare-ga nani-o tsukaimashita-ka?
who-Nom what-Acc used-Q
'Who used what?'

d. *Non-reversible scrambled:*
Nani-o dare-ga tsukaimashita-ka?
what-Acc who-Nom used-Q

e. *Primes for (4.4) pictures:*
Dare-ga nani-ni norimashita-ka?
who-Acc what-on rode-Q?
'Who rode what?'

The primes for the additional 6 pictures were all in the canonical word order as shown in (4.5e) because the purpose of the inclusion of these items was to observe whether the primes might facilitate the correct choice of particles in picture description. The prime questions for the filler pictures were also all in the canonical order. For about half of the filler pictures, however, not all of the NPs in the prime questions were replaced with *wh*-NPs. For instance, if the picture depicted a situation in which a girl is swimming in the pool, the prime question for the picture was "The girl swam where?" In this case, the participants had a choice of describing the picture with the subject, "The girl swam in the pool.", or without the subject, "Swam in the pool.", because ellipses of NPs are common in Japanese when the NP is known from the context.

The primes were digitally recorded by a native speaker of Japanese. For those prime questions that had canonical/scrambled pairs, special care was taken in recording so that the two *wh*-NPs in each question would receive similar intonation and stress.

The prime questions and the pictures in the reversible and non-reversible condition were split into two counterbalanced lists. That is, if a picture was preceded by a prime question in the canonical order in one list, the same picture was preceded by a prime question in the scrambled order in the other list, and vice versa. The additional 6 items and the filler items were common across the two item lists. Each list consisted of 8 each of reversible canonical/scrambled and non-reversible canonical/scrambled sentences, the 6 additional items, and 34 filler items, for a total of 72 items. In an effort to alleviate the influence of the item presentation order, two pseudo-randomized versions were created for each item list. The primes and the pictures were inserted into Microsoft PowerPoint slides.

4.7.3 Procedures

4.7.3.1 Task 5

Similarly to Task 1, the participants were instructed to read the English sentence first and to fill in the parentheses with Japanese particles so that the completed Japanese sentence would have the same meaning as the preceding English sentence. The only difference from Task 1 was that the participants were instructed to write “X” in the parentheses when they felt that the particle obligatorily had to be omitted. This direction was to accommodate to (4.4f) above in which “6 kilograms” cannot be followed by any overt particle. The use of the topic marker *–wa* was discouraged for this task.

4.7.3.2 Task 6

Prior to the task, the participants were given a word list and were instructed to study the words in the list. The word list included all of the vocabulary items in the target pictures as well as about half of the vocabulary items that were in the filler pictures. Glosses in English were provided next to each Japanese word. In addition to checking the meaning of the words, the NNS participants were also asked to pronounce each word aloud, while the NS participants were only asked to study the words in the list. As with the previous syntactic persistence task, the participants were also asked to study pictures in a folder for about 5 minutes for the recognition task. The folder included all of the filler pictures but none of the target pictures. The participants were told that the following task would be a recognition task, and that they would be asked to indicate whether the pictures they would encounter during the task were among those in the folder that they had just studied.

The presentation of the primes (sound files) and the pictures was carried out on a laptop computer with Microsoft PowerPoint. In each trial, a sound icon and a picture were presented. The participants were instructed to listen to the sound file (prime question) by clicking the sound icon and to respond to the question based on the information provided in the picture. They were instructed to listen to the question again if they missed it. The participants were instructed that they could use either all the words or only the necessary words in the picture to respond to the question.[39] They were, however, told that the response would have to be in a full sentence and that a response such as *"Pool desu."* ('It is pool.') would not be acceptable – the verb in the picture had to be placed at the end of their responses. After responding to the question, they were asked to indicate whether they had seen the picture in the folder by saying "yes" or "no" aloud.

[39] This instruction was given for the purpose of encouraging the participants to listen to the question primes carefully and answer the questions, instead of simply describing the pictures using all the words in the pictures.

There were four practice items prior to the actual trial. The experimenter went through the practice items with the participant to make sure that he/she could perform the task as intended. The participants' responses were recorded with a digital recorder and were transcribed for analysis.

4.7.4 Scoring and analysis

4.7.4.1 Task 5

The "lax" scoring in Task 1 was adopted for the scoring of the reversible canonical/scrambled and non-reversible canonical/scrambled items in Task 5. Therefore, marking an accusative NP with the *–ni* particle was considered correct. For the reversible canonical items and the additional 6 items (those in 4.4), the "strict" scoring in Task 1 was applied, because the correct choice of particles in these sentences was of interest. Responses were given the score of 1 if they were correct, and 0 if they were incorrect. Responses were judged correct only if both of the particles in a given sentence were correct.

The accuracy data for the reversible canonical/scrambled and non-reversible canonical/scrambled items were analyzed by each participant group using linear mixed-effects modeling in R. The fixed effects were Reversibility and Word order, and the random effects were Subjects and Items.

4.7.4.2 Task 6

The same method as in Task 4 was adopted for the scoring of the reversible canonical/scrambled and non-reversible canonical/scrambled items. When items were accidentally skipped by the participants, they were classified as "NA" and treated as missing data. There were 5 instances (out of 960 opportunities) of skipping for the NNS participants, and one instance of skipping (out of 512) for the NS participants. The proportions of the Canonical, Scrambled, and Other responses for each participant group and for each condition, after the removal of the "NA" items, are provided in Appendix 4D. The method of analysis using the linear mixed-effects modeling was also the same as that of Task 4.

For the reversible canonical items and the additional 6 items, the "strict" scoring in Task 1 was applied. In order to observe whether the primes in Task 6 might facilitate the correct use of particles, the "strict" scores of the reversible canonical items and the additional 6 items in Tasks 5 and 6 were compared by participant group using linear mixed-effects modeling. The data were treated as missing in cases when a) the participants skipped the trial or b) the participants' responses included only one of the two NPs. Between the reversible canonical and the additional 6 items, there were 13 items (out of 420) that were treated as missing in the NNS data. There were no missing data points in the NS data.

4.8 Results of the follow-up tasks

4.8.1 Task 5

Figure 4.4 presents the mean accuracy rates of the reversible and non-reversible sentences by the Adv and Int groups.

As observed in Figure 4.4, the mean accuracy rates of the reversible and non-reversible sentences among the Adv participants were quite high. An analysis using a linear mixed-effects model indicated that the main effect of Word order was significant (t = 2.389, p = 0.0173) while the main effect of Reversibility was not significant (t = 1.555, p = 0.1303). The interaction of the two fixed effects was not significant ($\chi^2(1)$ = 1.0499, p = 0.3055). When a linear mixed-effects model with the interaction term was used to analyze the data, the effect of Word order was significant for the reversible sentences (t = 2.413, p = 0.0162), but it was not significant for the non-reversible sentences (t = 0.965, p = 0.335). An analysis with a random slopes model (by-subject and by-item for the effect of Word order) indicated that the main effect of Word order was not significant (t = 1.66, p = 0.108).

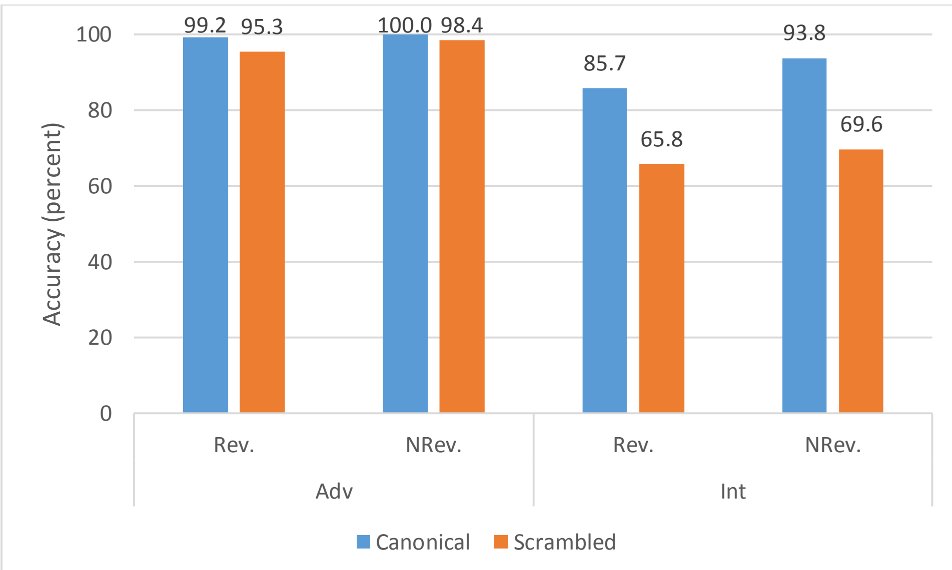


Figure 4.4: Mean accuracy rates in Task 5

For the Int group, the main effect of Word order was significant (t = 7.629, p < 0.0001), and the main effect of Reversibility was marginally significant (t = 2.045, p = 0.0415). The interaction of the two factors was not significant ($\chi^2(1)$ = 0.5637, p = 0.4528). An analysis using a random slopes model (by-subject for the effect of Word order) indicated that the main effect of Reversibility was not significant (t = 0.861, p = 0.4127). The main effect of Word order was still significant, but the effect was smaller (t = 2.707, p = 0.0206).

The results above seem to mirror those in Task 1 (with the “lax” scoring) in that the participants in the lower proficiency group were significantly less accurate with the scrambled sentences than the canonical sentences.

4.8.2 Task 6

The utterance rates of scrambled sentences by condition and by participant group are provided in Figure 4.5.

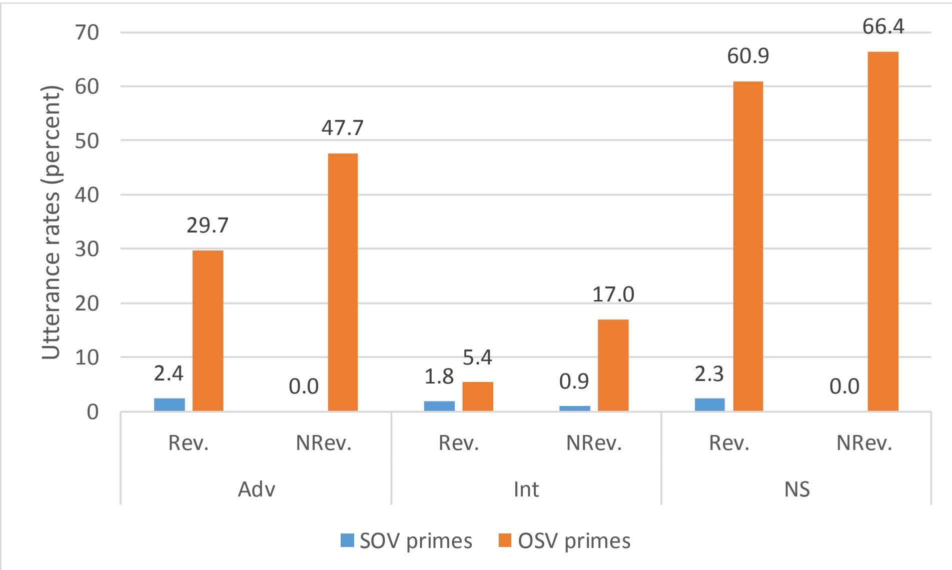


Figure 4.5: Utterance rates of scrambled sentences in Task 6

For the Adv group, the main effect of Word order was significant (t = 14.048, p < 0.0001). The main effect of Reversibility was also significant (t = 2.628, p = 0.0134). The interaction of the two fixed effects was significant (t = 3.874, p = 0.0001). The effect of Word order was significant for both the reversible and non-reversible sentences (reversible: t = 7.318, p < 0.0001; non-reversible: t = 12.837, p < 0.0001). When the same dataset was analyzed with a random slopes model (by-subject and by-item for the effect of Word order), the effect of Word order was smaller although it was still significant for both sentence types (reversible: t = 2.732; non-reversible: t = 4.769).

For the Int group, the main effect of Word order was significant (t = 4.562, p < 0.0001). The main effect of Reversibility was also significant (t = 2.498, p = 0.0129). The interaction of the two factors was significant (t = 2.927, p = 0.0036). The effect of Word order was significant for the non-reversible sentences ($t = 5.323$, $p < 0.0001$) while it was not significant for the reversible sentences ($t = 1.191$, $p = 0.2344$).

For the NS group, the main effect of Word order was significant ($t = 21.115$, $p < 0.0001$) while the main effect of Reversibility was not significant ($t = 0.526$, $p = 0.603$). The interaction of the two factors was not significant ($\chi^2(1) = 1.7536$, $p = 0.1854$). An analysis using a random slopes model (by-subject for the effect of Reversibility and Word order AND by-item for the effect of Word order) indicated that the main effect of Word order was still significant, but the effect was smaller ($t = 6.856$, $p = 0.0039$).

The results above indicated that, overall, the prime questions in OSV word order elicited the production of OSV sentences significantly more frequently than the questions in SOV order. The only exception was the reversible condition of the Int group in which the effect of Word order did not reach significance. The effect of the OSV questions was more robust with the non-reversible sentences than the reversible ones for the two NNS groups, but the reversibility did not seem to influence the syntactic persistence effect as much for the NS group.

4.8.3 Correct use of particles in Tasks 5 and 6

For the purpose of observing whether the primes in Task 6 might facilitate the correct use of particles, the “strict” scores of the reversible canonical items and the additional 6 items in Tasks 5 and 6 were compared by participant group using linear mixed-effects modeling. The fixed effect was Task, and the random effects were Subject and Verb. Verb represents the actual verbs used in the sentence items. Verb was entered as a random effect instead of Item because the sentence items between the two tasks were not exactly identical but the same verbs were used for the two tasks

Figure 4.6 presents the mean accuracy rates of the reversible canonical sentences by the two proficiency groups.

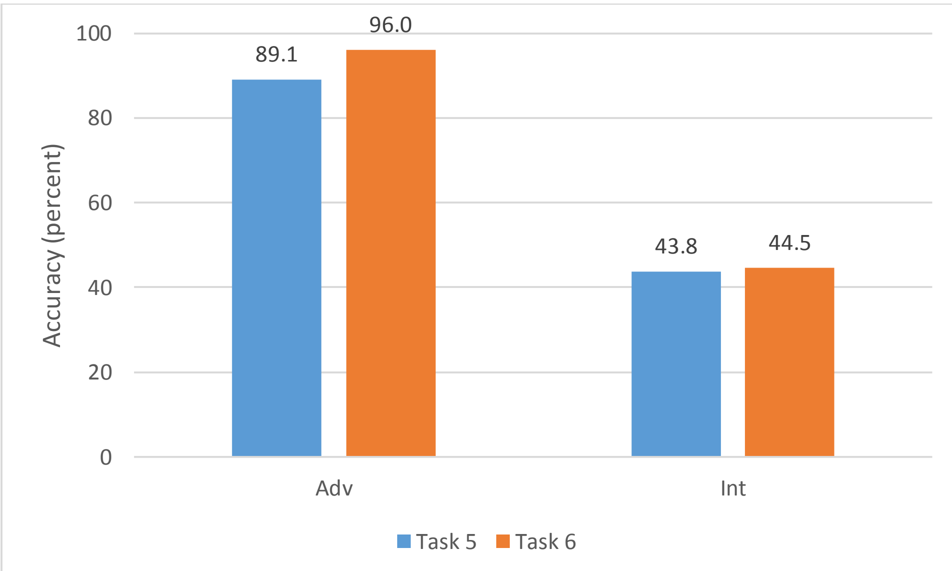


Figure 4.6: Mean accuracy rates for the reversible canonical sentences in Tasks 5 and 6

For the Adv group, the mean accuracy rate was significantly higher for Task 6 than Task 5 ($t = 2.251$, $p = 0.0253$). However, an analysis with a random slopes model (by-subject and by-verb) indicated that the effect of Task was not significant ($t = 1.152$, $p = 0.283$). For the Int group, the difference in mean accuracy rates in the two tasks was not significant ($t = 0.218$, $p = 0.8275$).

Figure 4.7 shows the mean accuracy rates for each of the additional 6 items by the two groups. The figures seem to indicate a more accurate production of particles in Task 6.

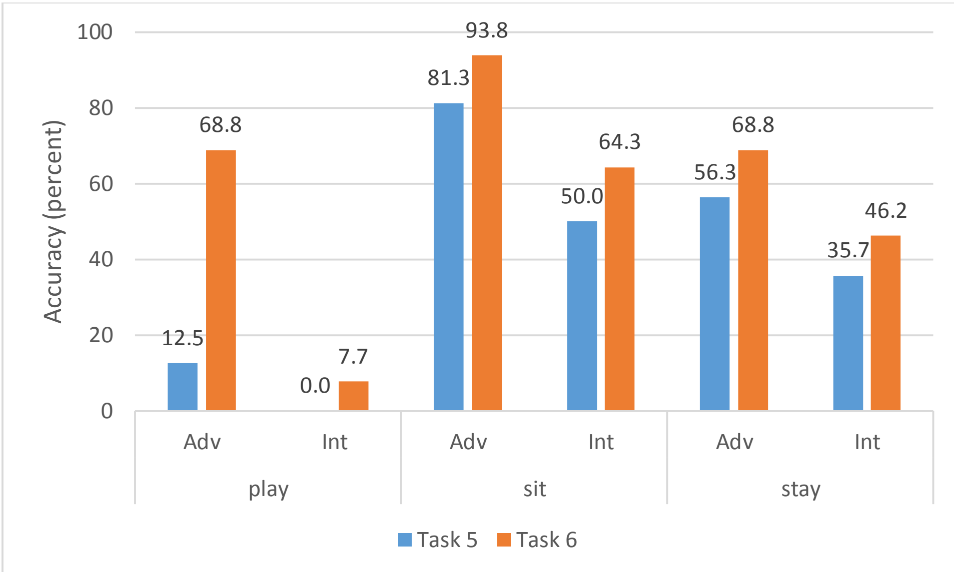


Figure 4.7 (Part 1): Mean accuracy rates for each of the additional 6 items in Tasks 5 and 6

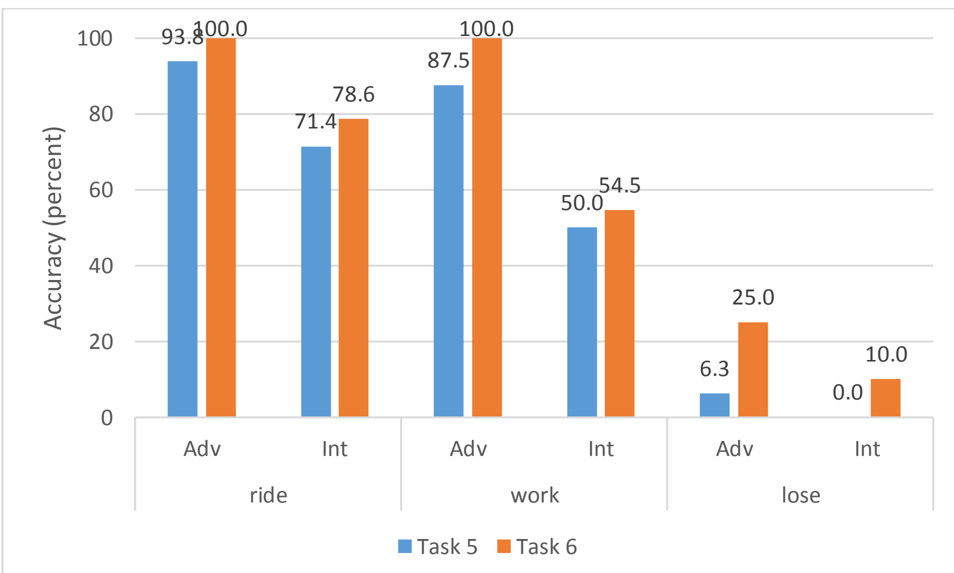


Figure 4.7 (Part 2): Mean accuracy rates for each of the additional 6 items in Tasks 5 and 6

While the particles in these 6 items were of different types, the data from the items were compiled for the purpose of comparing their mean accuracy rates between the two tasks. Figure 4.8 shows the mean accuracy rates for the six items by proficiency group and by task. An analysis with a linear mixed-effects model indicated that the mean accuracy rate was significantly higher for Task 6 than Task 5 for the Adv group (t = 3.842, p = 0.0002). On the other hand, the difference in mean accuracy rates in the two tasks did not reach the level of significance for the Int group (t = 1.442, p = 0.1514).

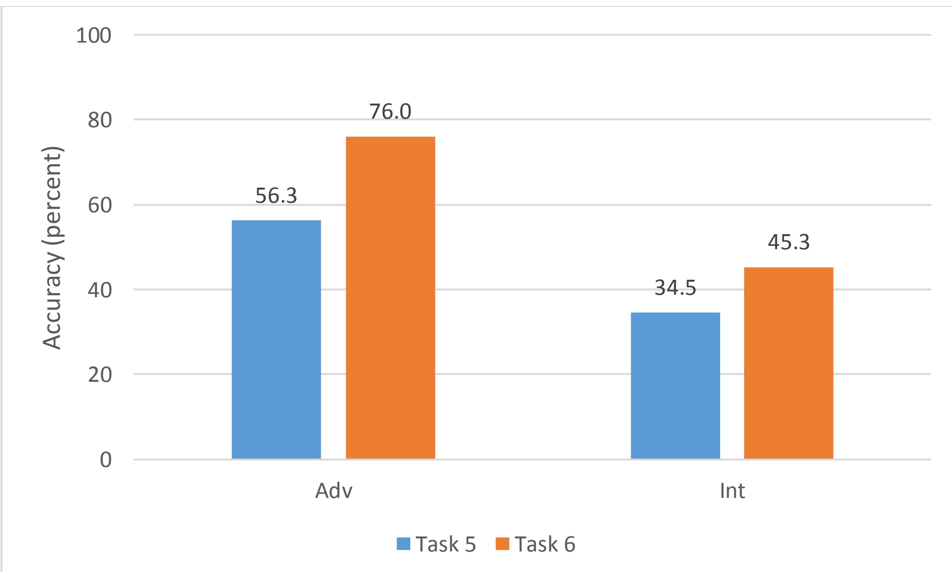


Figure 4.8: Mean accuracy rates for the additional 6 items in Tasks 5 and 6

To summarize the results above, the prime questions improved the choice of particles among the Adv participants. For the Int group, on the other hand, the "Q & A" format was not as effective in improving the choice of the particles.

4.8.4 Results summary of the follow-up tasks

Task 6 attempted to examine whether primes in a question format might facilitate the production of OSV sentences among L2 learners and NSs. Additionally, for the purpose of observing whether the question primes in Task 6 might facilitate the correct use of particles, the accuracy rates of the reversible canonical items and the additional 6 items in Tasks 5 and 6 were compared. The results indicated that the prime questions in OSV word order elicited the production of OSV sentences more frequently than the questions in SOV order. However, the production rates of OSV sentences were lower for the Int group compared with the Adv and NS groups, and the effect of OSV primes (as compared to SOV primes) did not reach significance for the reversible condition among the Int participants. The results also indicated that the OSV question primes elicited the production of OSV sentences more frequently for the non-reversible condition than the reversible condition among the NNS participants. That is, the interaction of Reversibility and Word order was significant for both the Adv and Int groups. On the other hand, reversibility of the items did not seem to influence the syntactic persistence effect as much for the NS group. As for the effect of the question primes on the use of postpositional particles, the method demonstrated a positive effect for the Adv group. For the Int participants, on the other hand, the prime questions were not as effective in improving their use of particles.

Overall, the "Q & A" syntactic persistence method did indicate some positive effects in the production of sentences in OSV order as well as the correct choice of particles among the L2 learners. However, the results, especially the ones regarding the production of OSV sentences, need to be interpreted with caution. Due to the limited availability of potential participants, those who already participated (or were going to participate) in the previous tasks were invited to participate in Tasks 5 and 6, and about half of the participants in those two tasks had already taken the debriefing interview for the previous tasks. Although care was taken to alleviate the effect of the debriefing interview (i.e., inviting the participants only if it had been more than two month since they took the debriefing interview), some of the individuals might have been quite aware of the purpose of the tasks (i.e., the OSV word order). Therefore, such a heightened awareness of the OSV structure might have increased the production of OSV sentences among the participants. In order to evaluate the effectiveness of the question primes more fully, a further examination with a different group of participants will be required.[40]

4.9 General discussion

While the results of Task 4 indicated a general absence of syntactic persistence of the OSV structure as well as low overall OSV production rates, Task 6 observed a significant effect for OSV prime questions and a more frequent overall production of OSV sentences. Since the primary difference between the two tasks was the type of primes (regular declarative OSV sentences in Task 4 and questions in OSV order in Task 6), the question primes in Task 6 might have contributed to the increase of OSV sentence production. Yamashita (2002), through analyses of various types of texts, suggests that most of the Japanese scrambling is motivated systematically by such features as "heaviness" or "referentiality" of the scrambled constituents,

[40] For the purpose of examining the effect of the debriefing interview, the utterance rates of OSV sentences among those who had already been debriefed when they took Task 6 and those who hadn't were compared by participant group using linear mixed-effects modeling. In this analysis, the data from the two reversibility conditions were collapsed, and the fixed effects were Word order and Debriefing. For the Adv group, the interaction of Word order and Debriefing was significant (t = 4.811, p < 0.0001), where the difference in utterance rates between the SOV/OSV prime conditions was larger among those who had not been debriefed. The utterance rate of OSV sentences in the OSV prime condition was also significantly higher among those who had not been debriefed (51.8% vs. 28.5%; t = 2.664, p = 0.0167). However, these differences did not reach the level of significance with an analysis using a random slopes model (by-subject for the effect of Word order). For the Int group, no statistically significant difference was observed between the two groups. For the NS group, the interaction of Word order and Debriefing was significant (t = 4.163, p < 0.0001), where the difference in utterance rates between two prime conditions was larger among those who had not been debriefed. The utterance rate of OSV sentences in the OSV prime condition was also significantly higher among those who had not been debriefed (71.9% vs. 50.0%; t = 3.297, p = 0.0033). In an analysis with a random slopes model (by-subject and by-item for the effect of Word order), the interaction of the two fixed effects was marginally significant (t = 2.118, p = 0.0495), and the difference in the OSV utterance rates between the two groups approached significance (t = 1.964, p = 0.067). Thus, for the Adv and NS groups, those participants who had not been debriefed tended to produce OSV sentences more frequently in Task 6 than those who had been debriefed.

and therefore, scrambled sentences “are not random variants of canonical order” (p. 617). Since scrambling itself does not seem to optimize comprehension, Yamashita further suggests that the motivation for scrambling may be production-based. One possibility, Yamashita argues, is that scrambling contributes to the efficiency of the speaker’s sentence production. The other possible production-based motivation for scrambling is that, due to the prominence or the accessibility of a particular constituent, the speaker might want to say that constituent first. In Task 4, the prime sentences were presented separately from the target pictures without any contexts. Thus, the task design might not have provided motivations (other than the prime sentences themselves) for the participants to use the scrambled word order to describe the pictures. In Task 6, on the other hand, the primes were direct questions about the target pictures. Therefore, it is speculated that the OSV question primes might have satisfied some necessary functional conditions for OSV sentences to be produced, perhaps either by the prominence of the fronted accusative *wh*-NP or by the “alignment between interlocutors” function of syntactic persistence (Pickering & Garrod, 2004) or both, resulting in increased overall production of OSV sentences.

Such a speculation might be supported by the production rates of OSV sentences by each participant group in Tasks 4 and 6. In Task 4, the only significant effect of syntactic persistence was observed in the reversible condition among the LI participants. The NS and ADV participants, who presumably had a better abstract syntactic representation of the “Object-Subject” word order, did not indicate the effect of the OSV primes and produced OSV sentences relatively infrequently. In Task 6, on the other hand, the production of OSV sentences was most frequent among the NS group, followed by the Adv group. The magnitude of the priming effect was also larger for the NS and Adv groups, while the Int group indicated a smaller priming effect and a lower production rate of OSV sentences. The difference between the two tasks may be explained if we assumed, as discussed above, that the production of scrambled sentences is not arbitrary, and that certain functional conditions need to be met for the production of scrambled sentences to take place. In Task 4, it might have been the case that the task design did not satisfy the conditions for the production of OSV sentences. Thus, the NS and ADV participants did not produce OSV sentences although they had the grammatical knowledge to do so, because the task did not provide intrinsic motivation for them to express the message in a scrambled word order. The canonical SOV word order is quantitatively far more frequent than the scrambled word order, and therefore, the subtle “encouragement” through syntactic persistence might not have been strong enough to cause OSV sentence production. The marginally significant effect of syntactic priming was observed in the reversible condition among the LI participants, perhaps because their knowledge of how to use the reversible verbs was not as established, which in turn might have made them more susceptible to syntactic persistence. In Task 6, on the other hand, the OSV prime questions might have provided a necessary cue for OSV sentence production, which perhaps motivated the NS and Adv participants to describe the subsequent target pictures in the OSV word order. The smaller priming effect as well as lower overall production rate of OSV sentences among the Int participants may suggest that they were less sensitive to the functional conditions that motivate the production of scrambled sentences.

The results of Task 6 indicated that the OSV question primes elicited the production of OSV sentences more frequently for the non-reversible condition than for the reversible condition among the NNS participants. This was perhaps due to the difference in the prime questions in the two conditions. In the reversible condition, because both the agent and patient in the target pictures were humans, the word *dare* (‘who’) was used to ask both the agent and patient. Thus, the only clue that the question primes were scrambled was provided by the case markers. In the non-reversible condition, on the other hand, the *wh*-NPs (*dare* 'who' vs. *nani* 'what') as well as the case markers signaled scrambling in the prime questions. This might have made the scrambling in the non-reversible prime sentences more salient, leading to a more robust syntactic persistence effect among the NNS groups. The NS data did not indicate a significant difference in the priming effects in the two conditions. It is probably because the NSs were fully able to integrate the case-marker information when they comprehended the reversible prime questions.

The results of Task 6 also indicated that, while the prime questions improved the use of particles for the Adv participants, they were not as effective for the Int participants. One possible cause of this difference is the amount of attention the participants were able to pay to the form of the prime questions when they comprehended them. VanPatten (2004) suggests that, due to working memory constraints and their attention to content words, L2 learners will be able to process input for form only after they are able to process it for meaning. Since processing of Japanese might have been more psychologically costly for the Int participants than the Adv counterparts, the Int participants might have been "too busy" comprehending the meanings of the prime questions to pay attention to their forms (i.e., particles). Another possible cause of the difference between the two groups is the hypotheses that the participants might have had about the particles. Due to their higher proficiency in Japanese, the Adv participants might have had multiple hypotheses about the particles to form the sentences correctly. Thus, when they heard the correct particles in the prime question, they might have been able to select the correct hypothesis about the particles using the prime question as a model. On the other hand, the Int participants might have had just one hypothesis about the correct particles or multiple hypotheses without the correct alternative. In such a case, they could not have selected the correct particles even with the assistance of the prime questions. For example, in (4.4a) 'Hanako played with the toy.', the correct particle for "with" is the *–de* particle because the toy is the "tool" with which Hanako played. On the other hand, if someone plays with a friend, the *–to* particle is used because "friend" is an animate entity. The Adv participants, due to their greater experience with Japanese, might have activated both *–de* and *–to* as alternatives to form (4.4a) correctly (although *–to* might have been more activated than *–de* initially). Thus, when they heard the *–de* particle in the prime question, they might have been able to choose the correct alternative to form the sentence. For the Int participants, on the other hand, they might have failed to recognize a toy as a tool, and thus, the *–de* particle might have been much less activated than the *–to* particle. Therefore, they might not have been able to choose the correct particle even after they heard *–de* in the prime question.

While Task 6 was more effective than Task 4 in eliciting the production of OSV sentences, the task as is still may not be an effective instructional tool. Task 6 was not very effective in eliciting OSV sentences and improving the use of postpositional particles among the Int participants. In addition, the production rate of OSV sentences among the Adv participants after the OSV prime questions was still low (29.7% and 47.7% for the reversible and non-reversible conditions, respectively). One potential improvement may be made by combining the methods of Tasks 6 and 2, in which the prime question precedes the target picture, and the target picture also specifies the item to start the description with. In such a method, the prime questions may provide a context for OSV sentences while the target pictures "force" learners to start the description with the object, potentially leading to a more frequent production of OSV sentences. The task can be easily adopted in classroom instruction, as the instructor can orally ask the prime question while pointing at the item with which the learners should start the description.

Another potentially effective method is the combination of explicit grammar instruction (word order and particles) and the method described above. As reviewed above, Shin (2008) observed that explicit instruction combined with the syntactic persistence task was more beneficial in the short-term learning of L2 structures. Moreover, explicit instruction, involving illustration of rules and use of metalinguistic knowledge, has generally been considered effective in facilitating the speed of L2 learning (Norris & Ortega, 2000), and some L2 instructional models have integrated explicit instruction of grammar as their essential component (e.g., focus on form: Long, 1991; form-focused instruction: Spada, 1997; processing instruction; VanPatten, 2002). Thus, it is expected that explicit instruction, combined with implicit learning through syntactic persistence, may further facilitate the learning of word order and postpositional particles. It would be of pedagogical importance to examine whether such a method might contribute to the acquisition of the scrambling word order and the use of particles, and if so, whether it might be effective in the long run.

CHAPTER 5: SUMMARY AND CONCLUSION

The present chapter summarizes the results of the three articles (Chapters 2-4) and provides provisional answers to the research questions presented in Chapter 1. It also discusses pedagogical implications of the present study and suggests some future research directions.

5.1 Summaries of the three articles

5.1.1 Chapter 2

Chapter 2 reported the results of Tasks 1 and 2. Task 1 (fill-in-the-blank) attempted to assess L2 Japanese learners' grammatical knowledge as the basis for processing scrambled sentences. Task 2 (picture description) attempted to assess their production performance of scrambled sentences.

The mean accuracy data of Task 1 based on the "lax" scoring (in which the marking of the accusative NPs with particles other than *–o* and not with *-ga* was considered correct) indicated that the participants across proficiency groups were generally less accurate in their responses with scrambled sentences than canonical sentences, and that the participants in the LI group were significantly less accurate with the scrambled sentences than the other two groups. The analysis of responses by individual participants also indicated that there was a rather large variability in L2 learners' grammatical knowledge, even within the proficiency groups. There were two participants in the ADV group who formed sentences with the OSV word order in fewer than 50% of the opportunities. On the other hand, about half of the LI participants consistently demonstrated their grammatical knowledge of scrambling. The "strict" analysis of the mean accuracy data (in which the marking of the accusative NPs with *–o* was considered the only correct answer) indicated that accusative marking in the reversible sentences was quite inaccurate (less than 50%, even for the canonical sentences) among the HI and LI participants.

In Task 2, the analysis of the mean accuracy data based on the “lax” scoring indicated that all three of the NNS groups were less accurate in producing scrambled sentences than canonical ones. The analysis also indicated that, while the interaction of Reversibility and Word order was not significant for the LI group, it was significant for the ADV and HI groups, suggesting that the participants in these two groups experienced more difficulty in producing reversible scrambled sentences than non-reversible scrambled sentences. Similarly to Task 1, the data in Task 2 indicated that there were substantial individual differences in the NNS participants’ production performance of OSV sentences within the proficiency groups. For instance, while 11 out of 20 LI participants produced non-reversible scrambled sentences using the correct particles 50% or more of the time, there were 7 participants in the same group who produced sentences in an SOV order 5 or more times in describing the non-reversible scrambled pictures. The reaction time data indicated that not only the NNS participants but the NS participants took longer to start producing scrambled sentences, suggesting an additional psychological cost in producing OSV sentences.

The results of the two tasks were further compared to examine the relationship of grammatical knowledge and production performance of OSV sentences, and the analysis suggested that the relationship might be different between reversible and non-reversible sentences. For the reversible sentences, while the participants’ performance on the two tasks mostly corresponded, there were cases in which the participants who demonstrated a grammatical understanding of scrambling did not necessarily perform well in production. There were more errors producing scrambled sentences, especially among the ADV and HI groups, and at least some of the errors seem to have been caused by the overuse of the SOV template although there is a possibility that the participants' relative unfamiliarity with reversible verbs might have contributed to the errors. For the non-reversible sentences, on the other hand, the grammatical knowledge of scrambling and production matched very closely, and there was little evidence that the errors in the production of OSV sentences were caused by the overuse of the SOV template.

5.1.2 Chapter 3

Chapter 3 was aimed at investigating the comprehension processes of OSV sentences by L2 learners, and reported the results of two tasks: the pilot study (sentence correctness decision task) and Task 3 (self-paced reading task).

The results of the pilot study indicated that scrambled sentences caused longer reaction times and lower accuracy rates among NS and NNS participants, suggesting that L2 learners of Japanese experience additional psychological cost in reading and comprehending scrambled sentences as do NSs. However, a difference in the sensitivity to case-marker information was observed between NNSs and NSs. While it took significantly longer for the NSs to read and comprehend the reversible scrambled sentences than the canonical ones, the difference did not reach the level of significance for the NNS group, which suggests that the NNS participants might not have integrated the case-marker information as consistently as the NSs.

While the results of the pilot study indicated that both NNSs and NSs took longer to read and comprehend OSV sentences than SOV sentences, the results of Task 3 suggested that the processing of OSV sentences by L2 learners might be quite different from that of NSs. The NS participants indicated slowdowns at the second NP position when they read the OSV sentences, as also observed in Mazuka et al. (2002). On the other hand, the NNS groups, regardless of their proficiency level, did not indicate such slowdowns. Thus, the results of Task 3 were consistent with Clahsen and Felser's (2006) claim that L2 sentence processing may not involve a native-like syntactic processing. Although the data from the ADV participants did not indicate slowdowns at the second NP position, they indicated slowdowns at the "NP1 + 1" and the Verb positions when they read the reversible scrambled sentences, which were not observed in the data of the HI and LI groups. Therefore, it was suggested that the ADV participants might have integrated the case-marker information more consistently than the HI and LI participants.

5.1.3 Chapter 4

Chapter 4 attempted to examine whether the Japanese OSV word order might persist among L2 learners using syntactic persistence tasks. It was for the purpose of exploring the possibility of using syntactic persistence as a means to elicit correct production of OSV sentences among L2 learners. While the main task (Task 4) used regular SOV/OSV sentences as primes, the smaller-scale follow-up study (Task 6) used questions in SOV/OSV orders as primes. In addition, Task 6 attempted to examine whether postpositional particles might prime, using results of Task 5 (a fill-in-the-blank task) as baselines.

The results of Task 4 indicated that the method was quite ineffective in eliciting the production of OSV sentences. While the reversible condition among the LI participants observed a marginally significant effect of syntactic persistence, there were no other significant syntactic persistence effects observed in the data. Moreover, the overall production rates of OSV sentences in the task were quite low. Since the production of OSV sentences was very infrequent among the NS and ADV participants (who presumably had a better-established abstract syntactic representation of the “Object-Subject” word order), it was suggested that the design of the task might not have provided enough motivation for the participants to use the OSV sentences to describe the pictures.

The results of Task 6 were more positive; the prime questions in OSV word order elicited the production of OSV sentences more frequently than the questions in SOV order. However, the production rates of OSV sentences were lower for the Int group compared with the Adv and NS groups, and the effect of OSV primes, as compared to SOV primes, did not reach significance for the reversible condition among the Int participants. The comparison of the data from Tasks 5 and 6 indicated that the prime questions were effective in improving the use of postpositional particles for the Adv group, but they were not as effective for the Int group. As a potential improvement to use the task as an instructional method, combining the methods of Tasks 6 and 2 was suggested, in which the prime question precedes the target picture, and the target picture also specifies the item with which to start the description. It was also suggested that combining the syntactic persistence task with explicit instruction might also be an effective instructional method, as demonstrated in Shin (2008).

5.2 Answers to the research questions

Question 1: Do L2 learners experience difficulty in producing OSV sentences as compared to SOV sentences? If so, is the source of difficulty the lack of grammatical knowledge or a processing problem?

The results of Task 2 indicated that all three of the NNS groups were less accurate in producing OSV sentences as compared to SOV sentences. Thus, it may be concluded that the production of OSV sentences is more difficult for L2 learners, regardless of their proficiency levels. As for the source of difficulty in producing OSV sentences, the comparison of the data from Tasks 1 and 2 indicated that it was largely the lack of grammatical knowledge that caused the production difficulty.

As observed, however, the relationship of grammatical knowledge and production performance was different between the reversible and non-reversible sentences. For the reversible scrambled sentences, there were some discrepancies between the participants' grammatical knowledge and production performance, and there was evidence that the production errors of such sentences were partly due to the overuse of the SOV template. The results also indicated that the overuse of the SOV template might persist even at a higher proficiency level. For the non-reversible scrambled sentences, on the other hand, the NNS participants' grammatical knowledge and production performance matched very closely, and there was little evidence that a processing problem such as the overuse of the SOV template caused the production difficulty.

Question 2: Is the comprehension process of OSV sentences different between L2 learners and NSs?

The results of Task 3 indicated that the NS participants slowed down at the second NP position when they read the OSV sentences, suggesting that this was where the reanalysis of the sentence structure took place. On the other hand, there was no clear evidence that any of the NNS groups slowed down at the second NP position. Therefore, the results suggested that the comprehension process of OSV sentences may be different between L2 learners and NSs, providing evidence for Clahsen and Felser's (2006) Shallow Structure Hypothesis.

As suggested in Chapter 3, however, there may be at least two potential reasons why Task 3 failed to observe native-like processing of OSV sentences among the NNS participants. The first potential reason is that the sentence stimuli in Task 3 might have been too complex for the NNS participants to engage in native-like syntactic processing. The second reason is that the NNS participants in the present study might not have been proficient enough in the target language. Future investigations are necessary to clarify whether such reasons might have actually contributed to the absence of native-like processing among the L2 learners.

Question 3: Is syntactic persistence effective in facilitating the production of OSV sentences? Does it improve the use of postpositional particles among L2 learners?

While the results of Task 4 indicated a general absence of syntactic persistence of the OSV structure as well as the low overall OSV production rates, Task 6 revealed a significant effect of the OSV prime questions and more frequent overall production of OSV sentences. Thus, it may be concluded that, with a proper method, a syntactic persistence task can be effective in facilitating the production of OSV sentences. In Task 4, the prime sentences were presented separately from the target pictures without any contexts. In Task 6, on the other hand, the primes were direct questions about the target pictures. Therefore, it was suggested that the OSV question primes might have satisfied some necessary functional conditions for OSV sentences to be produced, perhaps either through the prominence of the fronted accusative *wh*-NP or through the "alignment between interlocutors" function of syntactic persistence (Pickering & Garrod, 2004) or both, resulting in increased syntactic persistence effects and overall production of OSV sentences. However, the production rates of OSV sentences were lower for the Int group compared with the Adv and NS groups in Task 6.

The results of Task 6 also indicated that, while the prime questions improved the use of particles for the Adv participants, they were not as effective for the Int participants. Additional studies are necessary to investigate whether alternative methods suggested (i.e., combining the methods of Tasks 6 and 2, integration of explicit grammar instruction) may be effective in improving the use of postpositional particles as well as in facilitating the production of OSV sentences among L2 learners of Japanese with lower proficiency.

Question 4: What is the relationship between the grammatical knowledge/processing of the scrambled sentences and L2 learners' general proficiency in Japanese? Are more proficient learners better at processing OSV sentences?

The results of Tasks 1 and 2 indicated that the ADV and HI participants performed significantly better than the LI participants, suggesting that, on average, there is a positive relationship between L2 learners' overall proficiency and their grammatical knowledge/production performance of scrambling. As observed, however, there were substantial individual differences in the participants' knowledge and production performance even within the proficiency groups. Moreover, the ADV and HI participants, who generally demonstrated a good grammatical understanding of scrambling, still seemed to experience some difficulty in producing the reversible scrambled sentences, possibly due to interference by the overuse of the canonical template.

While there was a positive relationship between the NNS participants' proficiency and their production of OSV sentences, the results of Task 3 indicated that the processing of the OSV sentences by the three NNS groups were similar, and even the ADV participants did not seem to engage in a native-like syntactic processing when they read the scrambled sentences. However, the ADV participants did indicate slowdowns at the "NP1 + 1" and the Verb positions when they read the reversible OSV sentences, which may be taken as evidence that they identified scrambling by using the case-marker information appropriately but with a delay compared with NSs. Such slowdowns in the reversible OSV sentences were not observed for the HI and LI groups.

Question 5: What are the effects of "reversibility" on L2 learners' processing of OSV sentences? Do L2 learners integrate the animacy of NPs during their processing of scrambled sentences?

The stimuli used in the present study were strictly separate between the reversible and the non-reversible conditions, and therefore, the differences observed in the two conditions need to be interpreted with caution. However, the results of the series of tasks suggested that the "inanimate-animate" order of the NPs made scrambling in the non-reversible OSV sentences more salient, facilitating identification and processing of such sentences among the NNS participants. For instance, Task 2 observed evidence that the production errors in reversible OSV sentences were partly due to the overuse of the SOV template, but there was little evidence that such a processing strategy caused errors in the production of non-reversible OSV sentences. The results of the pilot study indicated that the scrambling effect in terms of the reaction times was more robust for the non-reversible sentences than the reversible ones, which seems to suggest that scrambling in the non-reversible OSV sentences was identified more consistently when the NNS participants read and comprehended such sentences.

By the same token, the comparison of the two conditions also suggested that the NNS participants might have experienced greater difficulty in processing the reversible OSV sentences, in which the animacy contrast of the NPs was absent, and the only information signaling the scrambling was case marking. In the pilot study, while it took significantly longer for the NSs to read and comprehend the reversible scrambled sentences than the canonical ones, the difference did not reach the level of significance for the NNS group, suggesting the possibility that some NNS participants overlooked the case markers and did not notice scrambling when they read the reversible OSV sentences. In the self-paced reading task (Task 3), the reading time data of the HI and LI groups did not indicate evidence that identification of scrambling took place when they read the reversible OSV sentences. This again seems to suggest that they did not consistently integrate the case-marker information to identify the grammatical relations in the sentences.

5.3 Pedagogical implications and future research directions

One recurring observation in the series of tasks in the present study was that at least some of the NNS participants had difficulty in using case markers for theta-role assignment in scrambled sentences. While the responses during the debriefing interview indicated that the difficulty might have come from the lack of knowledge that scrambling is permissible in Japanese, it is speculated that the lack of knowledge about the typical function of each particle (e.g., *–ga* for nominative, *–o* for accusative) might have also contributed to the difficulty. There were participants who performed well in completing scrambled sentences in Task 1 even though they expressed in the debriefing interview that they had not known, prior to the participation to the study, that the OSV word order is allowed in Japanese. These participants clearly knew the general function of *–ga* and *–o*, and used that knowledge accordingly to complete the task. On the other hand, there were participants who frequently used "*–ga –o*" to complete the non-reversible scrambled sentences in Task 1 although the way in which the task was set up made it clear that the order of NPs was scrambled in those items. There were also participants who used "double object" (i.e., *–ni –o* and/or *–o –ni*) to complete the non-reversible scrambled sentences. It is speculated that those participants might not have had a good understanding of the typical functions of the case markers. Given that it is case markers that ultimately determine the grammatical relations in a clause in Japanese, effective instructional methods for the acquisition of case markers are called for.

Explicit instruction and practice on scrambling may serve that purpose, since processing of scrambled sentences requires that L2 learners be aware of the functions of case markers (and other postpositional particles) instead of relying on the canonical template. For production, the method suggested in Chapter 4 (i.e., question primes followed by picture description in which the picture specifies the first NP to be used) may be worth implementing. Another possibly effective method is to provide context prior to picture description, because such context may provide motivation for production of scrambled sentences by fulfilling the "referenciality" function of scrambling. For comprehension, such tasks as identification of the agent in a sentence and match/no-match judgments between a sentence and a picture may be effective. Tasks like those can be easily adopted into classroom activities and/or can be provided as self-study materials. Of course, future investigations with a well-controlled pre- and post-test design are necessary to examine the effectiveness of such tasks for the acquisition of the scrambling structure and postpositional particles.

Further investigations are also necessary to examine how L2 parser processes scrambled sentences. As pointed out earlier, although the results of the self-paced reading task (Task 3) indicated that the NNS participants' processing of the OSV sentences was different from that of NSs, the results could have been due to the complexity of the stimuli and/or the generally lower proficiency of the participants. Another possibility is that the L2 processing of Japanese scrambling may depend on whether scrambling is part of the native language (L1) of learners. Koda (1993) observed that L1 Korean learners of Japanese, whose L1 has postpositional case markers and allows scrambling, performed significantly better in their comprehension of OSV sentences than learners whose L1 was Chinese or English. Hara (2009), in a self-paced reading experiment of Japanese ditransitive sentences, observed that L1 Korean learners slowed down at the gap-implicating positions for the "short-scrambling" sentences (no intervening phrase between the scrambled accusative and the dative arguments), while L1 Chinese learners did not produce such slowdowns. These results suggest that shared grammatical features between L1 and L2 (or more generally, the overall proximity of the two languages) may facilitate the processing of the L2. It would be fruitful to compare the processing of scrambled sentences by different L1 groups, since such studies may have important implications as to how the sentence processing strategies established through the use of L1 may influence the processing of an L2.

APPENDIX 1A: GENERAL BACKGROUND INFORMATION OF THE NNS PARTICIPANTS

	Group	Gender	Age	Length of J. study	Highest J. class	Stay in Japan	Frequency of J. use
NNS01	HI	M	23	36	5th	0	6
NNS02	ADV	M	21	48	8th	0	7
NNS03	ADV	M	21	44	6th	4.5	5
NNS04	LI	M	27	20	4th	4	3
NNS05	LI	M	20	16	4th	1	7
NNS06	LI	F	19	24	4th	0	6
NNS07	LI	M	22	24	4th	0	6
NNS08	ADV	F	22	72	8th	1	6
NNS09	HI	M	22	120	4th	0	7
NNS10	ADV	M	29	36	8th	26	5
NNS11	ADV	F	21	48	6th	6	5
NNS12	HI	M	25	48	6th	11	6
NNS13	LI	M	23	24	6th	2	6
NNS14	HI	M	20	24	4th	0	6
NNS15	LI	M	20	24	4th	0	2
NNS16	LI	F	25	18	4th	0	5
NNS17	LI	M	20	20	4th	0	7
NNS18	ADV	M	20	20	4th	0	4
NNS19	HI	M	23	18	4th	0	5
NNS21	LI	M	21	18	4th	0	5
NNS22	HI	F	21	8	2nd	36	7
NNS23	LI	F	20	20	4th	0	4
NNS24	ADV	M	22	46	8th	0	5
NNS25	HI	M	23	48	8th	11	3
NNS26	ADV	M	36	48	8th	75	6
NNS27	ADV	M	32	36	6th	36	5
NNS28	LI	M	23	66	4th	0	5
NNS29	ADV	F	29	25	8th	63	6
NNS30	HI	F	29	60	8th	24	5
NNS31	HI	F	22	25	4th	0	3
NNS32	HI	M	25	48	6th	0	2
NNS33	HI	M	21	24	5th	1.5	2
NNS34	LI	M	27	32	6th	2	5
NNS35	LI	M	20	16	4th	1	5
NNS36	LI	F	23	24	4th	0	2
NNS37	HI	F	21	24	4th	0	3
NNS38	HI	M	22	16	4th	0	3
NNS39	LI	M	20	24	4th	0	5
NNS40	ADV	F	21	24	4th	0	5

	Group	Gender	Age	Length of J. study	Highest J. class	Stay in Japan	Frequency of J. use
NNS41	LI	M	22	60	4th	0.25	2
NNS42	ADV	M	27	36	6th	12	4
NNS43	HI	F	24	36	4th	0.5	4
NNS44	ADV	M	20	24	4th	0	6
NNS45	ADV	F	21	37	6th	11	7
NNS46	ADV	M	27	25	4th	5	5
NNS47	ADV	F	22	48	6th	1	4
NNS48	ADV	M	25	60	6th	0	5
NNS49	LI	M	35	16	4th	0	3
NNS50	LI	M	24	36	6th	0	3
NNS51	LI	F	21	19	4th	1.5	6
NNS52	ADV	M	24	48	8th	13	2
NNS53	ADV	F	27	120	4th	1	4
NNS54	ADV	M	64	416	4th	72	5
NNS55	LI	F	20	16	4th	0	2

Notes: Length of Japanese study and length of stay in Japan are in months. For the frequency of Japanese use, the participants indicated how often they used the Japanese language in the past six month using a 7-point Likert scale (1: not at all; 7: very frequently).

APPENDIX 1B: NNS PARTICIPANTS' SELF-REPORTED PROFICIENCY AND THEIR SPOT SCORES

	Group	L	S	R	W	Overall	SPOT B	SPOT A	SPOT A+B	Adjusted
NNS01	HI	5	5	4	4	4	45	35	80	73
NNS02	ADV	4	3	4	3	3	52	54	106	97
NNS03	ADV	5	4	4	3	4	54	47	101	92
NNS04	LI	5	4	2	2	5	21	17	38	30
NNS05	LI	4	3	3	3	4	27	24	51	45
NNS06	LI	3	3	2	2	3	38	24	62	54
NNS07	LI	3	3	3	3	3	40	32	72	62
NNS08	ADV	5	5	4	4	5	60	61	121	111
NNS09	HI	3	4	5	4	4	54	42	96	87
NNS10	ADV	5	4	4	3	5	53	50	103	94
NNS11	ADV	6	6	4	4	5	58	53	111	102
NNS12	HI	5	5	4	5	5	49	39	88	78
NNS13	LI	4	5	5	5	5	41	29	70	62
NNS14	HI	4	4	5	4	5	52	41	93	85
NNS15	LI	3	4	3	3	3	35	31	66	60
NNS16	LI	4	3	4	3	4	21	15	36	27
NNS17	LI	5	4	3	4	4	32	24	56	50
NNS18	ADV	5	5	4	4	5	53	48	101	91
NNS19	HI	3	4	2	3	4	46	41	87	79
NNS21	LI	4	3	4	4	3	39	31	70	63
NNS22	HI	5	3	3	2	4	49	44	93	83
NNS23	LI	4	4	5	4	5	37	33	70	61
NNS24	ADV	4	5	4	3	4	55	61	116	108
NNS25	HI	5	4	4	5	5	50	43	93	84
NNS26	ADV	6	6	6	5	6	59	60	119	109
NNS27	ADV	5	5	4	4	5	50	52	102	92
NNS28	LI	2	2	1	1	1	27	28	55	46
NNS29	ADV	5	5	4	3	5	59	60	119	107
NNS30	HI	5	4	5	4	4	49	41	90	83
NNS31	HI	4	3	4	3	3	46	40	86	77
NNS32	HI	4	3	3	2	3	46	44	90	83
NNS33	HI	5	4	3	3	4	46	40	86	77
NNS34	LI	4	4	3	2	4	35	19	54	48
NNS35	LI	5	6	4	5	5	36	20	56	47
NNS36	LI	3	4	4	3	3.5	26	29	55	46
NNS37	HI	4	3	4	3	4	42	41	83	74
NNS38	HI	2	2	3	3	3	45	42	87	78
NNS39	LI	4	3	4	4	3	43	29	72	63
NNS40	ADV	5	4	5	5	5	59	54	113	105

	Group	L	S	R	W	Overall	SPOT B	SPOT A	SPOT A+B	Adjusted
NNS41	LI	2	3	3	3	2.5	34	22	56	49
NNS42	ADV	5	4	3	4	4	59	53	112	103
NNS43	HI	5	4	5	4	5	47	36	83	77
NNS44	ADV	4	4	5	6	5	59	56	115	105
NNS45	ADV	5	4	4	4	4	57	51	108	99
NNS46	ADV	2	3	3	3	2	56	47	103	96
NNS47	ADV	5	5	4	3	5	55	60	115	105
NNS48	ADV	2	3	4	3	3	58	55	113	104
NNS49	LI	2	3	3	1	2	13	9	22	15
NNS50	LI	5	3	3	2	4	21	24	45	40
NNS51	LI	4	3	5	4	4	34	19	53	45
NNS52	ADV	5	4	5	3	5	58	56	114	105
NNS53	ADV	5	4	4	3	4	52	55	107	98
NNS54	ADV	5	4	5	4	5	51	50	101	92
NNS55	LI	4	3	3	2	4	29	26	55	49

Notes: L, S, R, and W stand for listening, speaking, reading, and writing, respectively. For the four skills and the overall proficiency, the participants indicated self-rated proficiencies on 7-point Likert scales (1: beginner; 7: native-like).

APPENDIX 2A: ITEMS FOR TASK 1

Note: Below are sentences in canonical word order. The scrambled sentences were made by reversing the order of the nominative NP and the accusative NP. For the locative sentences, the order of the nominative NP and the locative NP was switched to make the scrambled sentences.

Reversible sentences:

フランス人が京子を見た。	The French person saw Kyoko.
次郎がルームメートを見た。	Jiro saw his roommate.
和子が後輩をいじめた。	Kazuko bullied her junior.
女の子が博をいじめた。	The girl bullied Hiroshi.
課長が誠をしかった。	The section leader scolded Makoto.
内田が部下をしかった。	Uchida scolded his subordinate.
由美がボーイフレンドを起こした。	Yumi woke up her boyfriend.
男の人が直子を起こした。	The man woke up Naoko.
おばさんがロバートをたたいた。	The middle-aged woman hit Robert.
学が彼女をたたいた。	Manabu hit his girlfriend.
加藤がダンサーをほめた。	Kato praised the dancer.
おじいさんがメアリーをほめた。	The elderly man praised Mary.
イタリア人が山口を待った。	The Italian person waited for Yamaguchi.
進が先輩を待った。	Susumu waited for his senior.
文子が母親を呼んだ。	Fumiko called out to her mother.
部長が中川を呼んだ。	The general manager called out to Nakagawa.

Non-reversible sentences:

女の人がレポートを書いた。	The woman wrote the report.
渡辺が手紙を書いた。	Watanabe wrote the letter.
吉田がファッション雑誌を読んだ。	Yoshida read the fashion magazine.
サラリーマンが新聞を読んだ。	The salaryman read the newspaper.
男の子がピザを食べた。	The boy ate pizza.
聡が朝ごはんを食べた。	Satoshi ate breakfast.
久美子がジュースを飲んだ。	Kumiko drank juice.
ウェイターがビールを飲んだ。	The waiter drank beer.
かわいい猫がドアを開けた。	The cute cat opened the door.
美香が大きい箱を開けた。	Mika opened the large box.
徹が車を洗った。	Toru washed the car.
高校生が食器を洗った。	The high school student washed the dishes.
コックがフライパンを持った。	The cook held the frying pan.
田村がパソコンを持った。	Tamura held the computer.
上田が赤い着物を着た。	Ueda put on the red kimono.
大工が黒いTシャツを着た。	The carpenter put on the black t-shirt.

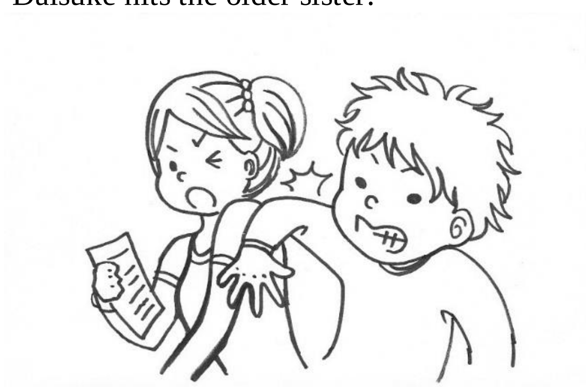
Locative sentences:

一郎が大学で勉強した。	Ichiro studied at the university.
花子が図書館で勉強した。	Hanako studied in the library.
剛が学校で運動した。	Tsuyoshi exercised at school
舞がジムで運動した。	Mai exercised in the gym.
謙二が家で遊んだ。	Kenji played in the house.
彩子が遊園地で遊んだ。	Ayako played in the amusement park.
彩子が海で泳いだ。	Ayako swam in the sea.
謙二がプールで泳いだ。	Kenji swam in the pool.
舞がボストンで死んだ。	Mai died in Boston.
剛がニューヨークで死んだ。	Tsuyoshi died in New York.
花子が寮で寝た。	Hanako slept in the dormitory.
一郎がアパートで寝た。	Ichiro slept in the apartment.

APPENDIX 2B: PICTURES USED FOR TASKS 2, 4, AND 6

<u>Pictures for the reversible condition:</u>

警官が洋子を見る。 ‘The policeman sees Yoko.’ 	田中が泥棒を見る。 ‘Tanaka sees the thief.’ 
和夫が弟をいじめる。 ‘Kazuo bullies the younger brother.’ 	お兄さんが友美をいじめる。 ‘The older brother bullies Tomomi.’ 
お母さんが謙二を起こす。 ‘The mother wakes up Kenji.’ 	理恵が赤ちゃんを起こす。 ‘Rie wakes up the baby.’ 

達也が小学生をしかる。 'Tatsuya scolds the elementary school student.' 	先生が小林をしかる。 'The teacher scolds Kobayashi.' 
上司が山田をほめる。 'The boss praises Yamada.' 	幸子が子どもをほめる。 'Sachiko praises the child.' 
大介がお姉さんをたたく。 'Daisuke hits the older sister.' 	夫が道子をたたく。 'The husband hits Michiko.' 

お父さんが恵子を待つ。 ‘The father waits for Keiko.’ 	健太が友達を待つ。 ‘Kenta waits for the friend.’ 
鈴木がウェイトレスを呼ぶ。 ‘Suzuki calls out to the waitress.’ 	妹が太郎を呼ぶ。 ‘The younger sister calls out to Taro.’ 

Pictures for the non-reversible condition:

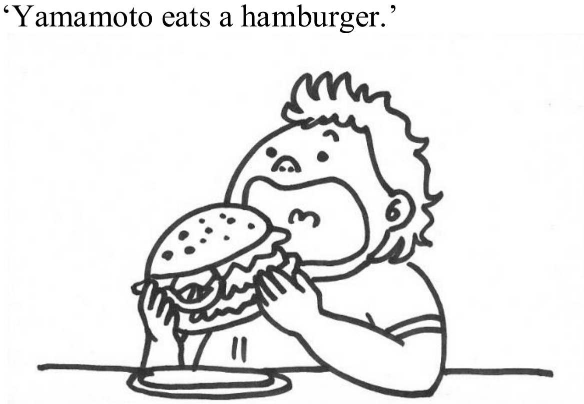
高木がメールを書く。
'Takagi writes an email.'

おじさんが手紙を書く。
'The middle-aged man writes a letter.'

医者が本を読む。
'The medical doctor reads a book.'

哲也が新聞を読む。
'Tetsuya reads a newspaper.'

山本がハンバーガーを食べる。
'Yamamoto eats a hamburger.'

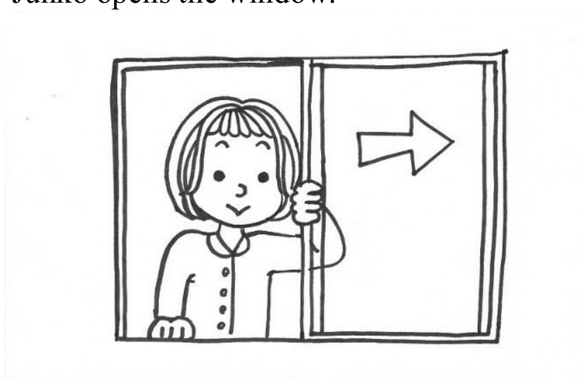


アメリカ人がステーキを食べる。
'The American person eats steak.'

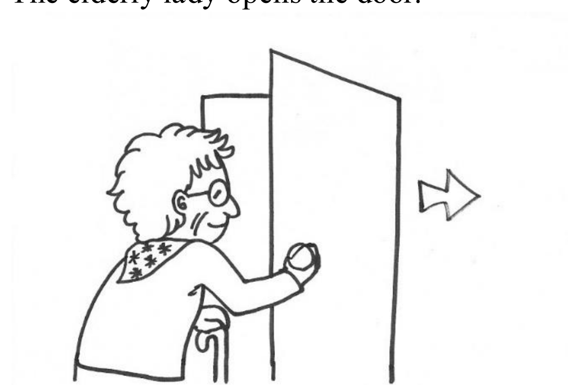
大学生がコーヒーを飲む。
'The college student drinks coffee.'

佐藤が薬を飲む。
'Sato drinks (takes) medicine.'

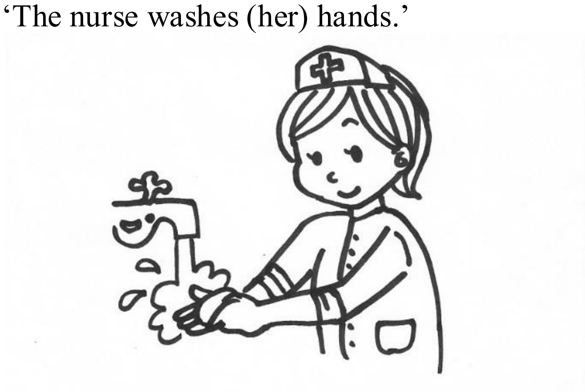
順子が窓を開ける。
'Junko opens the window.'



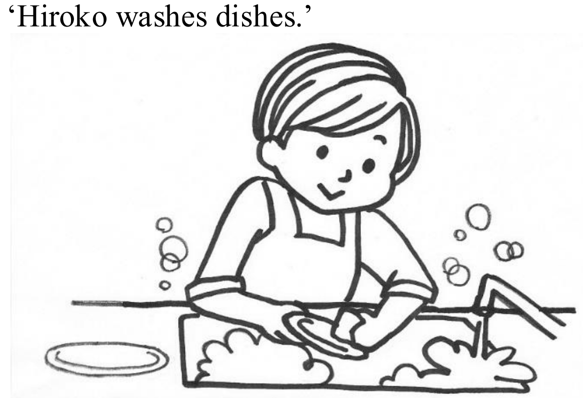
おばあさんがドアを開ける。
'The elderly lady opens the door.'



看護婦が手を洗う。
'The nurse washes (her) hands.'



弘子が皿を洗う。
'Hiroko washes dishes.'



木村が箱を持つ。 ‘Kimura holds the box.’ 	パイロットがかばんを持つ。 ‘The pilot holds the bag.’ 
イギリス人がセーターを着る。 ‘The British person puts on a sweater.’ 	洋平がシャツを着る。 ‘Yohei puts on a shirt.’ 

APPENDIX 3A: ITEMS FOR TASK 3

Note: Below are sentences in canonical word order. The scrambled sentences were made by switching the order of the nominative NP and the accusative NP.

Reversible sentences:
日曜日に達也が大きい銀行でお母さんを三十分待ったらしい。
It seems that Tatsuya, in the large bank, waited for his mother for thirty minutes on Sunday.

月曜日にお兄さんがにぎやかな空港で順子を三時間待ったようだ。
It appears that her older brother, in the busy airport, waited for Junko for three hours on Monday.

火曜日に友美が静かなコンビニで弟を何度も呼んだそうだ。
It is said that Tomomi, in the quiet convenience store, called out to her younger brother many times on Tuesday.

水曜日に社長がきれいなオフィスで和夫を二回呼んだらしい。
It seems that the president, in the pretty office, called out to Kazuo twice on Wednesday.

木曜日に大介が立派な公園で女の子を長々といじめたようだ。
It appears that Daisuke, in the splendid park, bullied the girl lengthily on Thursday.

金曜日に中学生が新しい体育館で花子をしつこくいじめたそうだ。
It is said that the junior high student, in the new gymnasium, bullied Hanako persistently on Friday.

土曜日に哲也が古い図書館で彼女を突然起こしたらしい。
It seems that Tetsuya, in an old library, woke his girlfriend up abruptly on Saturday.

先週サラリーマンが小さい映画館で太郎をいきなり起こしたようだ。
It appears that the salaryman, in a small movie theater, woke Taro up all of a sudden last week.

日曜日に鈴木が小さいレストランで高校生を長々としかったそうだ。
It is said that Suzuki, in a small restaurant, scolded the high school student lengthily on Sunday.

月曜日に上司が古い喫茶店で山本をしつこくしかったらしい。
It seems that the boss, in an old café, scolded Yamamoto persistently on Monday.

火曜日に道子が立派な会議室で部下を三時間ほめたようだ。
It appears that Michiko, in the splendid conference room, praised her subordinate for three hours on Tuesday.

水曜日におじいさんが新しい病院で洋子を三十分ほめたそうだ。
It is said that the elderly man, in the new hospital, praised Yoko for thirty minutes on Wednesday.

木曜日に渡辺がきれいなデパートで後輩をいきなりたたいたらしい。
It seems that Watanabe, in the pretty department store, hit her junior all of a sudden on Thursday.

金曜日に友達がにぎやかなスーパーで小林を突然たたいたようだ。
It appears that the friend, in the busy supermarket, hit Kobayashi abruptly on Friday.

先週久美子が静かな教会で有名人を二回見たそうだ。
It is said that Kumiko, in the quiet church, saw the celebrity twice last week.

土曜日に先生が大きい教室で洋平を何度も見たらしい。
It seems that the teacher, in the large classroom, saw Yohei many times on Saturday.

<u>Non-reversible sentences:</u>
先週医者が新しいオフィスでメールをゆっくり書いたようだ。
It appears that the doctor, in the new office, wrote the email slowly last week.

木曜日に健太が立派な図書館でレポートを急いで書いたそうだ。
It is said that Kenta, in the splendid library, wrote the report hastily on Thursday.

金曜日に弁護士が古い会議室で新聞を四十分読んだらしい。
It seems that the lawyer, in the old conference room, read the newspaper for forty minutes on Friday.

土曜日に弘子がにぎやかなコンビニで雑誌を二時間読んだようだ。
It appears that Hiroko, in the busy convenience store, read the magazine for two hours on Saturday.

日曜日におばさんが静かな銀行でクッキーを少し食べたそうだ。
It is said that the middle-aged lady, in the quiet bank, ate cookies a little on Sunday.

月曜日に中村が大きいレストランでハンバーガーをたくさん食べたらしい。
It seems that Nakamura, in the large restaurant, ate hamburgers a lot on Monday.

火曜日に中国人がきれいな映画館でコーヒーを少し飲んだようだ。
It appears that the Chinese person, in the pretty movie theater, drank coffee a little on Tuesday.

水曜日に三郎が小さい体育館でジュースをたくさん飲んだそうだ。
It is said that Saburo, in the small gymnasium, drank juice a lot on Wednesday.

先週おじさんが小さい病院で窓をあわてて開けたらしい。
It seems that the middle-aged man, in the small hospital, opened the window in a rush last week.

木曜日に幸子が静かな教室でプレゼントを丁寧に開けたようだ。
It appears that Sachiko, in the quiet classroom, opened the present carefully on Thursday.

金曜日に父親が大きい公園で車を丁寧に洗ったそうだ。
It is said that the father, in the large park, washed the car carefully on Friday.

土曜日に加藤がきれいな喫茶店で食器をあわてて洗ったらしい。
It seems that Kato, in the pretty café, washed the tableware in a rush on Saturday.

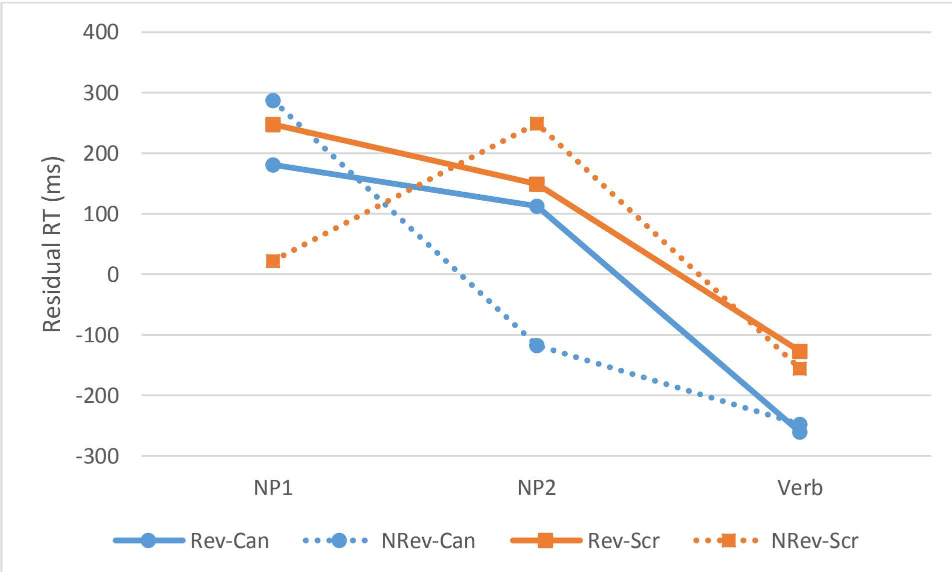
日曜日におばあさんがにぎやかなデパートで荷物を四十分持ったようだ。
It appears that the elderly lady, in the busy department store, held the luggage for forty minutes on Sunday.

月曜日に次郎が古い空港でスーツケースを二時間持ったそうだ。
It is said that Jiro, in the old airport, held the suitcase for two hours on Monday.

火曜日に大学生が立派な教会でジャケットをゆっくり着たらしい。
It seems that the college student, in the splendid church, put on the jacket slowly on Tuesday.

水曜日に礼子が新しいスーパーでセーターを急いで着たようだ。
It appears that Reiko, in the new supermarket, put on the sweater hastily on Wednesday.

APPENDIX 3B: MEAN RESIDUAL READING TIMES IN TASK 3 AMONG THE 10 PARTICIPANTS WHO SCORED HIGHEST IN SPOT



APPENDIX 4A: PRIME SENTENCES FOR TASK 4

Note: Below are sentences in canonical word order. The scrambled sentences were made by reversing the order of the nominative NP and the accusative NP.

<u>Reversible condition:</u>

美香が次郎をからかった。	Mika teased Jiro.
男の人がルームメートをからかった。	The man teased his roommate.
おじいさんが妻を殺した。	The elderly man killed his wife.
徹が伊藤を殺した。	Toru killed Ito.
渡辺が由美を探した。	Watanabe looked for Yumi.
男の客がウェーターを探した。	The male customer looked for the waiter.
犬がおばさんを助けた。	The dog helped the middle-aged lady.
直子が誠を助けた。	Naoko helped Makoto.
中村が京子を手伝った。	Nakamura assisted Kyoko.
サラリーマンがフランス人を手伝った。	The salaryman assisted the French person.
女の人がボーイフレンドを許した。	The woman forgave her boyfriend.
和子が聡を許した。	Kazuko forgave Satoshi.
一郎が吉田をけった。	Ichiro kicked Yoshida.
課長が部下をけった。	The section leader kicked his/her subordinate.
社長が弁護士を雇った。	The (company) president hired the lawyer.
井上が久美子を雇った。	Inoue hired Kumiko.

<u>Non-reversible condition:</u>

良平がパソコンを使った。	Ryohei used the computer.
佐々木が電話を使った。	Sasaki used the telephone.
和彦が辞書を買った。	Kazuhiko bought the dictionary.
池田がノートを買った。	Ikeda bought the notebook.
中国人が日本語を話した。	The Chinese person spoke Japanese.
高校生が英語を話した。	The high school student spoke English.
コックがケーキを作った。	The cook made the cake.
大工がいすを作った。	The carpenter made the chair.
男の子がズボンをはいた。	The boy put on the trousers.
中学生が靴をはいた。	The junior high student put on the shoes.
恵がピアノを弾いた。	Megumi played the piano.
直樹がギターを弾いた。	Naoki played the guitar.
斉藤が鉛筆を落とした。	Saito dropped the pencil.
愛子が消しゴムを落とした。	Aiko dropped the eraser.
猫が時計を壊した。	The cat broke the clock.
女の子がおもちゃを壊した。	The girl broke the toy.

APPENDIX 4B: PROPORTIONS (IN PERCENT) OF THE CANONICAL, SCRAMBLED, AND OTHER RESPONSES IN TASK 4

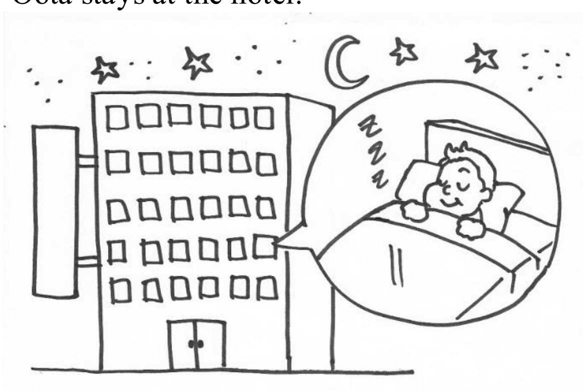
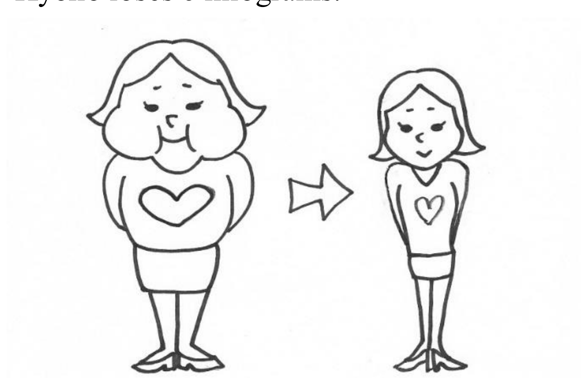
ADV		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	91.8	3.1	5.0
	Scrambled	93.5	3.2	3.3
Non-Reversible	Canonical	96.8	3.2	0
	Scrambled	97.5	2.5	0

HI		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	85.5	8.2	6.4
	Scrambled	90.7	6.5	2.8
Non-Reversible	Canonical	94.6	2.7	2.7
	Scrambled	99.1	0	0.9

LI		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	91.6	3.9	4.5
	Scrambled	83.4	7.9	8.6
Non-Reversible	Canonical	95.4	1.3	3.3
	Scrambled	96.2	2.6	1.3

NS		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	96.9	1.9	1.3
	Scrambled	96.3	1.3	2.5
Non-Reversible	Canonical	99.4	0.6	0
	Scrambled	98.7	1.3	0

APPENDIX 4C: THE SIX ADDITIONAL TARGET PICTURES USED IN TASK 6

女の子がおもちゃで遊ぶ。 'The girl plays with the toy.' 	男の子が椅子に座る。 'The boy sits on the chair.' 
大田がホテルに泊まる。 'Oota stays at the hotel.' 	男の人がバスに乗る。 'The man rides the bus.' 
徹が会社で働く。 'Tooru works at the company.' 	京子が 6 キロやせる。 'Kyoko loses 6 kilograms.' 

APPENDIX 4D: PROPORTIONS (IN PERCENT) OF THE CANONICAL, SCRAMBLED, AND OTHER RESPONSES IN TASK 6

Adv		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	96.8	2.4	0.8
	Scrambled	68.8	29.7	1.6
Non-Reversible	Canonical	99.2	0	0.8
	Scrambled	49.2	47.7	3.1

Int		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	90.1	1.8	8.1
	Scrambled	83.0	5.4	11.6
Non-Reversible	Canonical	89.1	0.9	10.0
	Scrambled	67.9	17.0	15.2

NS		Utterance Form		
Priming Condition		Canonical	Scrambled	Other
Reversible	Canonical	97.7	2.3	0
	Scrambled	38.3	60.9	0.8
Non-Reversible	Canonical	100.0	0	0
	Scrambled	33.6	66.4	0

REFERENCES

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Baldie, B. J. (1976). The acquisition of the passive voice. *Journal of Child Language*, *3*(3), 331-348.

Banno, E., Ohno, Y., Sakane, Y., Shinagawa, C., & Tokashiki, K. (1999). *Genki I & II: An integrated course in elementary Japanese*. Tokyo: The Japan Times.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, *11*(3), 245-299.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 931-949.

Berwick, R. C., & Fong, S. (1995). Madama Butterfly Redux: Parsing English and Japanese with a principles and parameters approach. In. R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing* (pp. 177-208). Hillsdale, NJ: Laurence Erlbaum.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-352). New York: John Wiley & Sons.

Bever, T. G., & McElree, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry*, *19*(1), 35-43.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18,* 355-387.

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition, 31,* 163-186.

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104,* 437-458.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129,* 177–192.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1-39.

Boyland, J. T., & Anderson, J. A. (1998). Evidence that syntactic priming is long-lasting. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the 20th annual conference of the Cognitive Science Society* (p. 1205). Hillsdale, NJ: Erlbaum.

Branigan, H. (2007). Syntactic priming. *Language and Linguistics Compass, 1,* 1-16.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75,* B13–B25.

Branigan, H. P., Pickering, M. J., Stewart, A. J., & McLean, J. F. (2000). Syntactic priming in spoken production: Linguistic and temporal interference. *Memory & Cognition, 28,* 1297-1302.

Chang, F., Bock, J. K., & Goldberg, A. (2003). Can thematic roles leave traces of their places? *Cognition, 90,* 29–49.

Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psychological Review, 113*, 234-272.

Chujo, K. (1983). Nihongo tanbun-no rikai katei – Bunrikai sutoratejii no sougo kankei [The interrelationship among strategies for sentence comprehension]. *Japanese Journal of Psychology, 54,* 250-256.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27,* 3-42; & Commentaries. *Applied Psycholinguistics, 27,* 43-105.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language, 49,* 214–230.

Desmet, T., & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language, 54,* 610–632.

Doi, Y., & Yoshioka, K. (1990). Jyosi no syuutoku ni okeru gengo unyoojoo no seiyaku: Pienemann-Johnston moderu no nihongo syuutoku kenkyuu e no ooyoo [Constraints on the acquisition of (Japanese) particles: Application of the Pienemann-Johnston model to acquisition of Japanese]. *Proceedings of 1st Conference on SLA and Teaching, 1*, 23-33.

Doi, Y., & Yoshioka, K. (1987). Which grammatical structure should be taught when? Implications of the Pienemann-Johnston Model to teaching Japanese as a foreign/second language. In *Proceedings of the Ninth Hawaii Association of Teachers of Japanese (HATJ) Conference on Japanese Language and Language Teaching*, 10-22.

Farmer, A. K. (1984). *Modularity in syntax: A study of Japanese and syntax.* Cambridge, MA: MIT Press.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164-203.

Ferreira, F., & Clifton Jr, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*(3), 348-368.

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying “that” is not saying “that” at all. *Journal of Memory and Language, 48,* 379–398.

Ferreira, V. S., & Bock, J. K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*(7-8), 1011-1029.

Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, *32*(6), 669-692.

Fletcher, C. R. (1985). The functional role of markedness in topic identification. *Text-Interdisciplinary Journal for the Study of Discourse*, *5*(1-2), 23-38.

Flett, S. J. (2006). *A comparison of syntactic representation and processing in first and second language production*. Unpublished doctoral dissertation. University of Edinburgh.

Ford-Niwa, J., & Kobayashi, N. (1999). SPOT: A test measuring “control” exercised by learners of Japanese. In K. Kanno (Ed.), *The acquisition of Japanese as a second language* (pp. 53-69). Amsterdam: John Benjamins.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124.

Forster, K. I., & Olbrei, I. (1972). Semantic heuristics and syntactic analysis. *Cognition*, *2*(3), 319-347.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 268-281). Oxford: Oxford University Press.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge: New York.

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

Goto, K. (1989). Tadoosi-bun ni okeru kakuzyosi “ga” “o” no siyoo to seigo-handan no hattatu. [Production and judgment of case particles “ga” and “o” in simple transitive sentences among preschool children]. *Tyoonoo Gengo-Gaku Kenkyuu*, *6*(1), 12-19.

Hakuta, K. (1982). Interaction between particles and word order in the comprehension and production of simple sentences in Japanese children. *Developmental Psychology, 18,* 62-76.

Hale, K. (1980). Remarks on Japanese phrase structure: Comments on the papers on Japanese syntax. *MIT Working Papers in Linguistics, 2,* 185-203.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English. Part 2. *Journal of Linguistics, 3,* 199-244.

Hara, M. (2009). Evidence for L2 syntactic gap-processing in Japanese scrambling sentences. *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)*, 54-63.

Harris, M. (1976). The influence of reversibility and truncation on the interpretation of the passive voice by young children. *British Journal of Psychology*, *67*(3), 419-427.

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214-238.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15,* 409-414.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*, B27-B39.

Hatasa, Y. A., Hatasa, K., & Makino, S. (2009). *Nakama I & II* (2nd ed.). Boston, MA: Heinle Cengage Learning.

Hayashibe, H. (1975). Word order and particles: A developmental study in Japanese. *Descriptive and Applied Linguistics*, *3,* 1-18.

Hayhurst, H. (1967). Some errors of young children in producing passive sentences. *Journal of Verbal Learning and Verbal Behavior*, *6*(4), 634-639.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22,* 369-397.

Hoshi, H. (1999). Passives. In N. Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 191-235). Malden, MA: Blackwell.

Inoue, A., & Fodor, J. (1995). Information-paced parsing in Japanese. In. R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing* (pp. 9-63). Hillsdale, NJ: Laurence Erlbaum.

Iwasaki, N. (2000). *Speaking Japanese: L1 and L2 grammatical encoding of case particles and adjectives/adjectival nouns*. Unpublished doctoral dissertation, The University of Arizona.

Iwasaki, N. (2003). L2 Acquisition of Japanese: Knowledge and use of case particles in SOV and OSV sentences. In S. Karimi (Ed.), *Word order and scrambling* (pp. 273-300). Malden, MA: Blackwell.

Iwatate, S. (1981). Nohongo-zi no syoki hatuwa ni okeru gozyun. [Word order in early utterances of Japanese children]. *Japanese Journal of Educational Psychology, 29*(2), 105-111.

Juffs, A., & Harrington, M. (1995). Parsing effects in second language processing: Subject and object asymmetries in wh-extractions. *Studies in Second Language Acquisition*, *17,* 483-516.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology, 11*(2), 228-238.

Kawaguchi, S. (2005). Argument structure and syntactic development in Japanese as a second language. In M. Pienemann (Ed.), *Cross-linguistic aspects of processing theory* (pp. 253-298). Philadelphia: John Benjamins.

Kilborn, K., & Ito, T. (1989). Sentence processing strategies in adult bilinguals. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 256-291). New York: Cambridge University Press.

Kim, J., Koizumi, M., Ikuta, N., Fukumitsu, Y., Kimura, N., Iwata, K., … Kawashima, R. (2009). Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics 22,* 151-166.

Kim, Y., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics, 29*(1), 149-154.

Kobayashi, N. (Ed.). (1997). *Development of SPOT (Simple Performance-Oriented Test) for the purpose of placing Japanese language students: Report (2)*. Japan: University of Tsukuba.

Kobayashi, N. (Ed.). (1998). *Development of SPOT (Simple Performance-Oriented Test) for the purpose of placing Japanese language students: Report (3)*. Japan: University of Tsukuba.

Koda, K. (1993). Transferred L1 strategies and L2 syntactic structures during L2 sentence comprehension. *Modern Language Journal, 77,* 490-500.

Kokuritsu Kokugo Kenkyujo. (1964). *Vocabulary and Chinese Characters in Ninety Magazines of Today,* vol 3. Tokyo: The National Language Research Institute.

Kondo, T., & Yamashita, H. (2011). Why speakers produce scrambled sentences: An analysis of a spoken language corpus in Japanese. In H. Yamashita, Y. Hirose & J. L. Packard (Eds.), *Processing and producing head-final structures* (pp. 195-215). Dordrecht: Springer.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming. Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

Kuno, S. (1973a). Japanese: A characteristic OV language. In W. Lehmann (Ed.), *Syntactic typology* (pp. 57-138). Austin, TX: University of Texas Press.

Kuno, S. (1973b). *The structure of Japanese language*. Cambridge, MA: MIT Press.

Kuroda, S.-Y. (1979). On Japanese passives. In G. Bedell, E. Kobayashi & M. Muraki (Eds.), *Exploration in linguistics: Papers in honor of Kazuko Inoue* (pp. 305-347). Tokyo: Kenkyuusha.

Kuroda, S.-Y. (1980). Bunpoo-no hikaku. [Comparative grammar]. In T. Kunihiro (Ed.), *Nichieigo hikaku kooza 2: Bunpoo* (pp. 23-62). Tokyo: Taishuukan.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package) http://CRAN.R-project.org/package=lmerTest

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics, 41,* 791–824.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, D. Coste, C. Kramsch & R. Dinsberg (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.

MacWhinney, B., & Bates, E. (Eds.). (1989). *The cross-linguistic study of sentence processing*. New York: Cambridge University Press.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27,* 53-78.

Mazuka, R., Itoh, K., & Kondo, T. (2002). Cost of scrambling in Japanese sentence processing. In M. Nakayama (Ed.), *Sentence processing in East Asian languages* (pp. 131-166). Stanford, CA: CSLI.

McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition, 28,* 179-207.

McDonough, K., & Kim, Y. (2009). Syntactic priming, type frequency, and EFL learners' production of wh-questions. *Modern Language Journal, 93*(3), 386-398.

McDonough, K. & Mackey, A. (2008). Syntactic priming and ESL question development. *Studies in Second Language Acquisition, 30,* 31-47.

Meijer, P. J. A., & Fox Tree, J. E. (2003). Building syntactic structures in speaking: A bilingual exploration. *Experimental Psychology, 50,* 184–195.

Mitsugi, S., & MacWhinney, B. (2010). Second language processing in Japanese scrambled sentences. In B. VanPatten & J. Jegerski (Eds.), *Research in second language processing and parsing* (pp. 159-175). Amsterdam: John Benjamins.

Miyagawa, S. (1989). *Syntax and semantics 22: Structure and case marking in Japanese*. San Diego, CA: Academic Press.

Miyahara, K. (1974). The acquisition of Japanese particles. *Journal of Child Language, 1*, 283-286.

Miyamoto, E. T., & Nakamura, M. (2005). Unscrambling some misconceptions: A comment on Koizumi and Tamaoka (2004). *Gengo Kenkyu*, *128*, 113-129.

Miyamoto, E. T., & Takahashi, S. (2002). Sources of difficulty in Japanese sentence processing. In M. Nakayama (Ed.), *Sentence processing in East Asian languages* (pp. 167-188). Stanford, CA: CSLI.

Miyamoto, E. T., & Takahashi, S. (2004). Filler-gap dependencies in the processing of scrambling in Japanese. *Language and Linguistics 5*(1), 153-166.

Mizumoto, G. (2009). Yooji no bunrikai hattatu ni oyobosu sadookioku yooryoo no eikyoo: Nihongoji ni okeru tan'itu koobun no rikai kara [The effect of working memory capacity on the development of children's sentence comprehension: The case of single argument sentences]. *Kyuusyuu Daidaku Gengogaku Ronjyuu, 30*, 1-27.

Morishita, M., Satoi, H., & Yokokawa, H. (2010). Verb lexical representation of Japanese EFL learners: Syntactic priming during language production. *Journal of the Japan Society for Speech Sciences, 11*, 29-43.

Moriyama, S. (2008). *Ninchi gengogaku kara mita nihongo kakujoshi no imi kouzoo to syuutoku: Nihongo kyooiku ni ikasu tameni* [A cognitive linguistic view of Japanese case particles: Their semantic constructions and acquisition for teaching Japanese]. Tokyo: Hitsuji-shoboo.

Muraoka, S., Tamaoka, K., & Miyaoka, Y. (2004). Kakimaze-bun no rikai ni okeru kaku-jyoshi no eikyoo [The effects of case markers on processing scrambled sentences]. *Technical Report of IEICE (Institute of Electronics, Information and Communication Engineers). Thought and language*, 37-42.

Murasugi, K. (2000). Bunpoo kakutoku: Idou gensyoo-o tyuusin tosite [The acquisition of grammar with special reference to movement]. *Academia: Literature and Language, 68*, 223-259. Nanzan University.

Murasugi, K., & Kawamura, T. (2004). On the acquisition of scrambling in Japanese. *Language and Linguistics 5*(1), 131-151.

Nakano, Y., Felser, C., & Clahsen, H. (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research, 31*(5), 531-571.

Nakayama, M. (1990). *Accessibility to the antecedents in Japanese sentence processing*. Unpublished manuscript, Ohio State University, Columbus.

Nakayama, M. (1991). *Japanese motion verbs and probe recognition.* Unpublished manuscript, Ohio State University, Columbus.

Nakayama, M. (1995). Scrambling and probe recognition. In. R. Mazuka, & N. Nagai (Eds.), *Japanese sentence processing* (pp. 257-273). Hillsdale, NJ: Laurence Erlbaum.

National Institute for Japanese Language & National Institute of Information and Communications Technology. (2004). *The Corpus of Spontaneous Japanese*.

Nemoto, N. (1999). Scrambling. In N. Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 121-153). Malden, MA: Blackwell.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, *50*(3), 417-528.

Otsu, Y. (1994a). Case-marking particles and phrase structure in early Japanese acquisition. In B. Lust, M. Suner & J. Whitman (Eds.), *Heads, projection, and learnability* (pp. 159-169). Hillsdale, NJ: Lawrence Erlbaum Associates.

Otsu, Y. (1994b). Early acquisition of scrambling in Japanese. In T. Hoekstra & B. D. Schwartz (Eds.), *Language acquisition studies in generative grammar: Papers in honor of Kenneth Wexler from the 1991 GLOW workshop* (pp. 253-264). Amsterdam: John Benjamins.

Park, B. J. (2007). *Syntactic persistence within and across languages in English and Korean L1 and L2 speakers*. Unpublished doctoral dissertation. The University of Arizona.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39,* 633–651.

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language, 46,* 586–605.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134,* 427-459.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27,* 169-226.

Pritchett, B. L. (1991). Head position and parsing ambiguity. *Journal of Psycholinguistic Research, 20*(3), 251-270.

Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago, IL: University of Chicago Press.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition, 47,* 59–87.

Rounds, P. L., & Kanagy, R. (1988). Acquiring linguistic cues to identify agent. *Studies in Second Language Acquisition, 20,* 509-42.

Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Saito, M., & Hoji, H. (1983). Weak cross-over and move α in Japanese. *Natural Language & Linguistic Theory 1,* 245-259.

Salamoura, A., & Williams, J. N. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual lexicon. *Applied Psycholinguistics, 28,* 627–660.

Sano, K. (1977). An experimental study on the acquisition of Japanese simple sentences and cleft sentences. *Descriptive and Applied Linguistics, 10,* 213-33.

Sasaki, Y. (1994). Paths of processing strategy transfers in learning Japanese and English as foreign languages. *Studies in Second Language Acquisition, 16,* 43-72.

Scheepers, C. (2003). Syntactic priming of relative clause attachments: Persistence of structural configuration in sentence production. *Cognition, 89,* 179-205.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, *56,* 153-171.

Shibatani, M. (1990). *The languages of Japan.* Cambridge: Cambridge University Press.

Shigenaga, Y. (2012). Processing of scrambled sentences by learners of Japanese as a second language. *Arizona Working Papers in Second Language Acquisition and Teaching, 19*, 79-103.

Shin, A. S. (2008). *Structural priming in bilingual language processing and second language learning*. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, *5*(3), 219-227.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12,* 229-265.

Spada, N. (1997). Form-focussed instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, *30*(2), 73-87.

Steinhauer, K., White, E., Cornell, S., Genesse, F., & White, L. (2006). The neural dynamics of second language acquisition: Evidence from event-related potentials. *Journal of Cognitive Neuroscience,* supplement, 99.

Sugimura, Y. (2002). Imeeji de oshieru nihongo no kakujoshi [Teaching of Japanese case particle through images]. *Gengo bunka kenkyuu soosyo, 1*, 39-55. Nagoya University.

Tamaoka, K. (2005). Tyuugokugo o bogo tosuru nihongo gakusyuusya niyoru seizyun/kakimaze gozyun no noodoobun to kanoobun no rikai [Comprehension of Japanese active and potential sentences with canonical and scrambled word orders by native Chinese speakers learning the Japanese language]. *Nihongo Bunpoo [Journal of Japanese Grammar], 5(*2), 92-109.

Tamaoka, K., Sakai, H., Kawahara, J., Miyaoka, Y., Lim, H., & Koizumi, M. (2005). Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research, 34*(3), 281-332.

Tanaka, M. (2008). *The representation of conceptual and syntactic information during sentence production*. Unpublished doctoral dissertation. University of Edinburgh.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules.* Cambridge, MA: MIT Press.

Trueswell, J., Tanenhaus, M., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285-318.

Turner, E. A., & Rommetveit, R. (1967). The acquisition of sentence voice and reversibility. *Child Development*, *38*(3), 649-660.

Ueno, M., & Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language*, *86*(2), 243-271.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, *52*(4), 755-803.

VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 1-31). Mahwah, NJ: Erlbaum.

Yamashita, H. (1997). The effects of word-order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research, 26*, 163-188.

Yamashita, H. (2002). Scrambled sentences in Japanese: Linguistic properties and motivations for production. *Interdisciplinary Journal for the Study of Discourse, 22*(4), 597-633.

Yamashita, H., & Chang, F. (2001). 'Long before short' preference in the production of a head-final language. *Cognition, 81*(2), B45-B55.

Yamashita, H., Chang, F., & Hirose, Y. (2002). Separating functions and positions: Evidence from structural priming in Japanese. Poster presented at the *14th Annual CUNY Conference on Human Sentence Processing*.