ABSTRACT

Title of Thesis:	INDIVIDUAL DIFFERENCES IN COMPREHENDING JAPANESE SCRAMBLED SENTENCES
	Yoko Eshita, Master of Arts, 2019
Thesis Directed By:	Dr. Steven J. Ross, Second Language Acquisition

This study's aim is to investigate further into the relationship between individual differences—working memory and sound recognition ability—and sentence processing of Japanese scrambled sentences for second language (L2) Japanese learners. L2 Japanese learners drawn from 3rd year college-level courses or above were tested on their listening comprehension accuracy in identifying case marking particles in canonical and scrambled sentences. Participants demonstrated a significant slowdown in reaction time and low accuracy rates for scrambled sentences compared with canonical sentences. In addition, even participants with high working memory and proficiency had difficulty in comprehending scrambled sentences and could not process case markings efficiently and accurately in a timed setting. This study is significant in that it is one of the first to examine the relationship between individual differences and comprehending Japanese case markings.

INDIVIDUAL DIFFERENCES IN COMPREHENDING JAPANESE SCRAMBLED SENTENCES

by

Yoko Eshita

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Arts
2019

Advisory Committee:
Professor Steven J. Ross, Chair
Professor Robert M. DeKeyser
Professor Nan Jiang

Acknowledgements

I would like to express my sincere gratitude to my advisor, Dr. Steven Ross. Without his continual guidance and kind support, it would not have been possible to complete my thesis. I am especially thankful for his patience throughout my thesis preparation and for helping me with the statistical analysis. My great appreciation also goes to my thesis committee members, Dr. Robert DeKeyser and Dr. Nan Jiang, for their suggestions and constructive feedback.

In addition, I would also like to thank Mori sensei, Oomori sensei, Hamano sensei, Tsujioka sensei, Marshall sensei, Sato sensei, and Kitamura sensei for helping me recruit participants. Lastly, I would like to thank the many contributors and supporters to the materials, the design, the data collection, and the writing process for this study.

Table of Contents

Acknowledgements ii
Table of Contents iii
List of Tables v
List of Figures vi
List of Abbreviations vii
Chapter 1: Introduction 1
Chapter 2: Review of the Literature 4
2.1 Japanese Case Marking Particles 4
2.1.1 Importance of Japanese Case Marking Particles 4
2.1.2 Native Speaker (NS) Japanese Processing Scrambled Sentences 5
2.1.3 English and Japanese Interlanguage Comprehension Strategy 7
2.1.4 Processing of Case Markings in Artificial Language: *Japlish* 8
2.1.5 L2 Japanese Processing Scrambled Sentences 10
2.2 Individual Differences 13
2.2.1 Working Memory 14
2.2.2 The Role of Working Memory (WM) in L2 Listening Comprehension 16
2.2.3 Foreign Language Aptitude: Sound Recognition Ability 17
2.2.4 The Role of Sound Recognition Ability in L2 Learning 18
2.3 Summary of the Review of the Literature 19
Chapter 3: Research Questions and Hypotheses 21
Chapter 4: Methods 24
4.1 Participants 24
4.2 Materials 25
4.2.1 Target Task 25
4.2.2 SPOT Test Version B 29
4.2.3 Listening Span Task 29
4.2.4 LLAMA-D 30
4.3 Procedures 31
4.4 Analysis 34
Chapter 5: Results 37
5.1 Accuracy of Scrambled Sentences 37
5.2 Individual Differences 38
5.2.1 Proficiency: SPOT Test 39
5.2.2 Working Memory 41
5.2.3 WM Recall 41
5.2.4 WM Total 43
5.2.5 WM Recall and SPOT 45
5.2.6 WM Total and SPOT 47
5.2.7 Interaction Between Variables 49
5.3 Reaction Time of Scrambled Sentences 51
Chapter 6: Discussion 53
6.1 Research Questions and Hypotheses 53
6.2 Discussion 56
6.3 Limitations 61

Chapter 7: Conclusions ... 63
Appendices ... 64
Appendix A: Target Task Sentences ... 64
Appendix B: Participant Background Questionnaire ... 67
Appendix C: Exit Survey ... 69
Bibliography ... 70

List of Tables

Table 1. *Self-Ratings of Japanese Language Skills* 25
Table 2. *Procedure of the Study* 35
Table 3. *Mean Accuracy and Accuracy Rate (in parenthesis) and Reaction Times (RTs in Milliseconds) for Each Sentence Type* 37
Table 4. *Paired Sample t-Test Among Sentence Types on Accuracy Rate* 38
Table 5. *The Odds Ratios and Probability of Being Accurate (in parenthesis) of the Interaction among Sentence Type and SPOT, WM Recall, and WM Total* 50
Table 6. *Paired Sample t-Test Among Sentence Types on Reaction Time* 51
Table 7. *Participants with High Accuracy on the Scrambled Sentences and their WM and SPOT Scores* 57

List of Figures

Figure 1. Procedure for DMDX.. 33
Figure 2. Histogram of Participants' Accuracy Rate among Sentence Types 37
Figure 3. Histogram of the Participants' Accuracy for SPOT Test Version B 39
Figure 4. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' SPOT Test Version B Scores 41
Figure 5. Histogram of Participants' Accuracy for WM Recall................................. 42
Figure 6. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Recall Scores.. 43
Figure 7. Histogram of Participants' Accuracy for WM Total 44
Figure 8. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Total Scores 45
Figure 9. Scatterplot of Participants' Accuracy in SPOT and WM Recall 46
Figure 10. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Recall and SPOT Test Version B Scores 47
Figure 11. Scatterplot of Participants' Accuracy on WM Total and SPOT (r = 0.25)48
Figure 12. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Total and SPOT Test Version B Scores 49
Figure 13. Bar Graph of the Probability of Being Accurate in Scrambled Sentences and Comparing its Interaction with SPOT, WM Recall, and WM Total............ 50

List of Abbreviations

ACC	Accusative case
DAT	Dative case
LOC	Locative case
NOM	Nominative case
PAST	Past tense

Chapter 1: Introduction

Japanese typology is a Subject-Object-Verb (SOV) language; however, its word order is quite flexible, with the exception that the verb must be placed at end of a clause. Therefore, both of the following sentences are grammatically correct and have the same meaning.

1) John-ga Mary-o mimashita.

-NOM -ACC see-PAST

'John saw Mary'

2) Mary-o John-ga mimashita.

-ACC -NOM see-PAST

This free word-order phenomenon is known as *scrambling*. Sentence 1 is a canonical SOV sentence, while Sentence 2 is a scrambled sentence. Japanese scrambling is made possible because of its case marking particles. These case marking particles are attached at the end of nouns to give clear identification for the role of the noun—they indicate the relationship of the preceding noun to the following noun or to the rest of the sentence (Yagi, 1992). As shown in the sentences above, the particle -*ga* is attached to the noun phrase, which serves as the subject of the sentence, and the particle -*o* marks the object and assigns an accusative role. Because of these case marking particles, even when the noun phrases are flipped in an Object-Subject-Verb (OSV) order, the meaning of the sentence remains virtually the same. Japanese speakers utilize these case marking participles along with word order when producing and comprehending sentences (Koda, 1993). Therefore, it is an important aspect of Japanese grammar.

Although the concept of scrambled sentences exists, the occurrence of scrambling is relatively low in written media. In writing, scrambled sentences occur approximately once in every 17 sentences in newspaper (Kuno, 1973), in less than 1 percent of sentences (19 scrambled sentences out of 2,635 sentences) in various texts types (Yamashita, 2002), and in about 6 percent of sentences (185 scrambled sentences out of 3,103 sentences) from a newspaper-based corpus (Orita, 2017). Even though scrambled sentences are used infrequently in writing, there still exists a chance that second language (L2) Japanese learners come across this structure while conversing with Japanese people; therefore, they must be proficient with it. The present study aims to investigate further the relationship between individual differences—Working Memory (WM) and sound recognition ability—and processing of case marking particles using Japanese scrambled sentences in L2 Japanese learners via a listening task.

The following chapter will first look more in depth at the use of Japanese case marking particles and how they are processed by native Japanese speakers and L2 Japanese learners. It will examine previous studies, especially those on L2 Japanese learners, on how the experiments were conducted and how it influenced shaping this project. Then it will look at why individual differences, such as WM and sound recognition ability, are important factors in L2 learning. Chapter 3 will address the research questions and the hypotheses of this study. Then, in Chapter 4, the methods of how this study will be carried out will be discussed. Following this will be the results of this study and a discussion of these results along with the limitations. The last chapter will be on the study's conclusion.

Chapter 2: Review of the Literature

2.1 Japanese Case Marking Particles

2.1.1 Importance of Japanese Case Marking Particles

Japanese case marking is often said to be one of the most difficult aspects for L2 Japanese learners, and many L2 learners struggle with comprehending them. One of the reasons behind the difficulty is that because there are only a few particles and each one has several meaning and functions (Yagi, 1992). For example, -*ga* can be used as a subject marker and a conjunction. Alternatively, -*ga* can be the subject marker while -*wa* is a topic marker in a single sentence even though both -*ga* and -*wa* are used to state the subject of the sentence (Kuno, 1973). Another reason that makes understanding Japanese case particles challenging is that case marking particles often cannot be directly translated into the learners' L1 (H. Suzuki & Toutanova, 2006). These properties of case marking particles result in added difficulty for L2 Japanese learners to grasp their proper meaning and usage.

L2 Japanese learners can often understand complex Japanese sentences; however, when it comes to producing them, they are prone to basic mistakes (Khan & Bryce, 2005). Even minor mistakes in the usage of case markings sometimes result in a different semantic interpretation from what the L2 Japanese learners intended to say (H. Suzuki & Toutanova, 2006). Moreover, case marking particle errors are found frequently in L2 Japanese learners' writing (Oyama, 2010) and among the three particles, -*ga* (subject marker), -*wa*, and -*o*, the L2 learners' accuracy rate in producing them was -*wa* > -*o* > -*ga* (subject marker) for L1 English L2 Japanese beginning- and intermediate-level learners (Yagi, 1992). A similar pattern emerged for L1 English L2 Japanese beginner-level learners in an oral interview (Doi & Yoshioka, 1988; Komori & Banno, 1988). Although some studies show different orders of acquisition (Mori, 2008; Yagi, 1996), they still indicate that the misuse of particles occurs commonly for L2 Japanese learners.

Another interesting aspect of Japanese case markings is that in conversation, many case markings are omitted except when the sentences are scrambled. In conversation, case markings are dropped because it is thought that it does not indicate the grammatical relations and is redundant (Hinds, 1982). Ellipsis—omission of linguistic elements that can be understood through context—often occurs in Japanese because it is a heavily context-dependent language (Kuno, 1978). In many cases, not only are case markings omitted but subjects or direct/indirect objects can also be omitted depending on the context, making word order even less reliable as a syntactic marking device (Koda, 1989). Therefore, L2 Japanese learners who have strict sentence structure in their first language (L1), such as in English, may struggle with the Japanese flexible word order. The understanding of case markings plays an important role in comprehending Japanese sentences especially because of these phenomena.

2.1.2 Native Speaker (NS) Japanese Processing Scrambled Sentences

Children's acquisition of scrambled sentences has been a topic of interest for many decades. In early studies, children until around 5 years old were said to not be able to process scrambled sentences correctly as they would comprehend the first noun being the subject (Hayashibe, 1975; K. Sano, 1977; S. Suzuki, 1977). However, recent studies show that children, between the ages of 3 to 5, are able to comprehend scrambled sentences when they are provided with context (Otsu, 1994; T. Sano, 1994) or when there are prosodic cues (Minai, Isobe, & Okabe, 2015). Basic linguistic knowledge on scrambling may be acquired at or before the age of three (Murasugi & Kawamura, 2005); however, without pragmatic or prosodic support, it becomes challenging for children to utilize their knowledge and results in difficulty with comprehending scrambled sentences (Minai et al., 2015). These studies suggest that NS Japanese children can utilize case marking particles in a certain context in order to comprehend scrambled sentences properly from a young age.

There are mixed results whether NS Japanese adults experience processing cost in comprehending scrambled sentences. One theory is that they do not—that reading scrambled sentences takes the same amount of time as reading canonical sentences. This may be because NS Japanese utilize case markings when reading sentences (Yamashita, 1997). Also, in a self-paced reading task, after each noun phrase, the reading time shortened suggesting that NS Japanese can infer the possible upcoming constituents based on the attached particles (Mitsugi & MacWhinney, 2010). Moreover, in a visual-world paradigm, NS Japanese were able to successfully utilize case marking particles to predict upcoming linguistic items in a listening task (Mitsugi & MacWhinney, 2016). These studies argue that case marking particles allow NS Japanese to successfully interpret scrambled sentences without any processing cost.

Other evidence, however, suggests that scrambled sentences will, in fact, incur processing costs. When reading scrambled sentences, an eye-tracker captured that participants tended to gaze longer and make more eye-movement regression when reading a scrambled sentence indicating that these sentences are processed online and cost more processing resources (Mazuka, Itoh, & Kondo, 2002). In addition, participants took a longer time to comprehend scrambled sentences when the accusative and dative cases are farther apart than when they are closer together (Miyamoto & Takahashi, 2004) suggesting that NS Japanese experience psycholinguistic measurable cost. Even for mono-clausal scrambled sentences, there was an indication of preverbal processing difficulty suggesting that there is, in fact, an online processing cost for even simple sentences (Witzel & Witzel, 2016). Moreover, when participants were presented with scrambled sentences, fMRI studies indicated that the left inferior frontal gyrus[1]—also known as Broca's area—became more activated than when participants were attempting to comprehend non-canonical sentences, suggesting that scrambled sentences impose more cognitive load (Kim et al., 2009). This research thus poses the opposite hypothesis—that scrambled sentences do cause a processing cost even for NS Japanese.

2.1.3 English and Japanese Interlanguage Comprehension Strategy

Although there has been abundant research on NS Japanese processing scrambled sentences, studies on L2 Japanese learners processing scrambled sentences are quite limited. There is, however, some research on what aspects of comprehension strategies transfer among L2 Japanese learners from English and whether they are perceived similarly or differently from NS Japanese. These studies have investigated the interpretation of case, animacy, and word order cues by NS Japanese and L2

[1] The left inferior frontal gyrus is responsible for syntactic processing across languages (Kim et al., 2009).

Japanese learners. In these studies, participants were requested to report the subject/agent after listening to Japanese word strings consisting of two nouns and a verb. These word strings are presented in various orders (i.e. noun-noun-verb, noun-verb-noun, etc.) to determine the strength of the cues. As for NS Japanese, the strength order is: case marker, followed by animacy, followed by word order (Ito, Tahara, & Park, 1989). These results support the fact that NS Japanese place much value on case marking. As for L2 Japanese, if there is no case marking in the word strings, intermediate- and advanced-level learners tended to favor animacy over word order (Sasaki, 1991). However, when case markings were added in the word strings, L2 Japanese speakers relied comparatively more on case markings but still less than on animacy (Sasaki, 1994). Moreover, L1 English L2 Japanese learners tended to comprehend noun-noun-verb word strings as if it were a SOV sentence, even if the first noun was an inanimate object, while the NS Japanese would comprehend the subject according to the case markings (Kilborn & Ito, 1989). These results suggest that shifting cue reliance from learners' L1 to L2 cannot be done rapidly and easily. In addition, L2 Japanese learners are likely to use the most frequent and available cues, such as SOV order, when comprehending Japanese words in order to complete the task (Matessa & Anderson, 2000). Therefore, comprehending strategies differ among NS and L2 Japanese learners, and L2 learners rely on the comprehension strategies of their L1.

2.1.4 Processing of Case Markings in Artificial Language: *Japlish*

Some studies have used a semi-artificial language called "*Japlish*" which uses English lexis while utilizing Japanese word order and case markings. *Japlish* has been used to control for frequency of exposure to word order and case marking particles. Although these studies (Grey, Williams, & Rebuschat, 2014, 2015; Williams, 2010; Williams & Kuribara, 2008) were mainly used to investigate incidental learning, they shed some light on how L2 or third language (L3) learners utilize case marking in sentence processing.

Participants who were not familiar with Japanese or any other languages that have case marking particles were asked to do a grammatical judgement test after having some training and exposure on *Japlish* sentences in both canonical and scrambled word order. 11 out of 25 participants (44%) who were previously trained and exposed to 114 *Japlish* sentences (Williams & Kuribara, 2008) and 7 out of 25 participants (28%) who were exposed to 388 *Japlish* sentences (Williams, 2010) had a strong preference for canonical word order, ignoring case markings. However, 13 of the 34 participants (38%) were able to correctly state the case marking rules for -*ga* and -*o*; 12 out of 34 participants (35%) for -*ni*; and 11 out of the 34 participants (32%) were able to provide a correct example of the functions of all of these case markers (Grey et al., 2014). Moreover, participants' accuracy rate increased (58% in the first quarter to 72% in the fourth quarter) throughout the exposure phrase showing that learning was taking place (Grey et al., 2015). These studies indicate that even with a small amount of exposure, some learners are able to grasp the rules and functions for case marking particles on their own accord. Since these studies on *Japlish* were using a semi-artificial language and participants were only exposed to this language for less than an hour before the test, results using natural language and testing learners who had been studying the language for a few years may differ.

2.1.5 L2 Japanese Processing Scrambled Sentences

Studies on the processing of scrambled sentences using Japanese among L2 learners are quite limited and their results are often contradictory. This section will describe each experiment in detail and form the framework for this paper.

Koda (1993) conducted an experiment on listening comprehension with canonical and scrambled sentences with and without particles. Participants consisted of beginning level L2 Japanese learners who were L1 English, Chinese, and Korean speakers with at least 3 years of previous non-Japanese L2 learning. These participants were asked to listen to canonical and scrambled Japanese sentences with and without case marking particles and then identify the subject/agent of the sentences. Accuracy in identifying the subject/agent was compared among different L1 backgrounds with regard to whether particles were present or absent. Results showed that participants, regardless of their L1 backgrounds, were able to perform with greater accuracy when the particles were present in both canonical and scrambled sentences. A non-significant trend towards better performance was noted in the L1 English and Chinese learners. These results suggest that Japanese sentence comprehension in adult L2 learners benefits from case marking.

In a reading task, Iwasaki (2003) examined different proficiency levels of L1 English L2 Japanese learners. Participants were asked to indicate whether the sentences presented were grammatically correct. These sentences included some with incorrect particles along with scrambled sentences. The participants' reaction time (RT) increased and their accuracy rate decreased when they read scrambled sentences, regardless of their proficiency. These results indicated that participants' knowledge of scrambled sentences was not as established compared with canonical sentences. Moreover, participants were asked to do a fill-in-the-blank task with the appropriate particles to measure their knowledge about case marking. Although there was a trend that as the participants' Japanese proficiency became higher, their knowledge about case marking particles increased in canonical sentences, this trend was not demonstrated for scrambled sentences.

In another study related to reading, Shigenaga (2012) studied the presence of slowdowns in L1 English speakers comprehending scrambled sentences using self-paced reading. Participants were asked to read the sentences, some of which had inanimate objects, then decide whether each sentence sounded correct or not. The results showed that there is a psychological cost in reading and comprehending scrambled sentences and that RT and error rate were higher when the sentence included an inanimate object. A particular subset of participants continuously comprehended the first noun as the subject/agent in all sentences indicating that these participants were reading all sentences in SOV structure, and thus not all the L2 Japanese learners had acquired the importance of case marking particles in Japanese sentences.

Other studies have demonstrated that L2 Japanese learners are able to use case markings properly. Mitsugi and MacWhinney (2010) tested L2 Japanese learners who were L1 English and Korean speakers using self-paced reading. The results showed that there was no significant difference in the reading time between canonical and scrambled sentences although a non-significant trend towards slower reading speed in the L1 Korean group was noted when reading scrambled sentences. This study concludes that there was no clear additional processing cost associated with comprehending scrambled sentences.

Similarly, Smith (2016) conducted an experiment with intermediate- and advanced-level L1 English L2 Japanese learners using self-paced reading followed by a comprehension question. All the participants were able to comprehend scrambled sentences without any errors. Smith (2016) suggested that "L2 Japanese learners demonstrate a high degree of accuracy with respect to both the production and comprehension of case marking, and that this is true even at relatively low proficiency levels" (p.118), which indicates that learners can comprehend case marking particles even from an early stage and utilize them effectively. It should be noted, however, that there were comparatively fewer sentences in Smith's study compared with the previous experiments.

In a visual-world paradigm, Mitsugi and MacWhinney (2016) conducted an eye-tracking experiment while listening to Japanese sentences. Participants were L1 English L2 Japanese learners who were enrolled in or had completed third- and fourth-year Japanese courses in the US. They were asked to listen to a Japanese sentence while looking at computer screen which projected four images. After listening to the sentence, participants were asked a comprehension question in English. Although L2 Japanese learners had a good grammatical knowledge about case marking particles in a grammar task and a high accuracy rate for the comprehension questions, they were not able to utilize the case markings fully to predict upcoming linguistic items. This may be because L2 learners may not be able to successfully predict oncoming information especially during online processing and may rely on other nonlinguistic information in order to account for their lack of processing speed.

With the limited number of studies with conflicting measurements and outcomes, further research is needed in the use of case marking particles in comprehending sentences among L2 Japanese. Moreover, since there are few studies looking into why some L2 Japanese learners incur a slowdown or have higher error rates when comprehending scrambled sentences, this study focuses on the relationship between the processing of Japanese case marking particles in a listening task and individual differences. The following part will examine in depth two individual differences, Working Memory (WM) and sound recognition ability, and how they play a positive role in L2 learning. It will also discuss the type of tasks that would be best suited for this study.

2.2 Individual Differences

Individual differences can play a role in L2 learning. Some learners can achieve high levels of L2 proficiency with relative ease, while others have a hard time even after the same amount of exposure (Miyake & Friedman, 1998). Since this study is aimed to investigate whether there is a relationship between individual differences, especially in WM and sound recognition ability, in processing case marking particles in Japanese sentence, the current section will focus on these individual differences and how they affect L2 learning.

2.2.1 Working Memory

One of the goals of this study is to examine the relationship between WM and the processing of case marking particles in scrambled sentences. WM is one of the most researched factors relating to individual differences in cognition (Martin & Ellis, 2012) since it is thought to be important in language comprehension (Daneman & Merikle, 1996; Harrington & Sawyer, 1992; Leeser, 2007; Miyake & Friedman, 1998; Waters & Caplan, 1996). One of the most influential WM models is Baddeley's Working Memory Model (Baddeley, 1986, 2003a, 2003b) which is compromised of the following components: "the phonological loop is said to be a component of the larger multicomponential WM system, which also includes a central executive (involved in allocating attention and regulating information processing), a visuospatial sketchpad (which, along with the phonological loop, is seen as a 'slave system' for storage of modality-specific information), and an episodic buffer (to integrate information from different modalities)" (Mackey & Sachs, 2012, p. 709). The phonological loop can further be divided into two subcomponents, "a temporary storage system which [holds] memory traces over a matter of seconds, during which they decayed, unless refreshed by the second component" (Baddeley, 2003b, p. 191). Since the phonological loop is an online process, if there are too many elements to process at any one time, those elements may fade before they can be reactivated and rehearsed (Gathercole, 2006) affecting the efficiency and the quality of language processing (Miyake & Friedman, 1998).

There have been a variety of methods to measure WM. The simple short-term storage capacity of WM is measured by methods such as digit span recall tasks, word-span tasks, non-word repetition tasks, and sentence repetition tasks; while a more complex WM, which demands more storage and processing function, utilizes alternative methods such as the Reading Span Test and the Listening Span Test (Miki, 2012). In a meta-analysis between various types of WM tasks and L2 outcomes, WM was positively associated with L2 processing and proficiency outcomes, with a population effect size (ρ) of around 0.225 (Linck, Osthus, Koeth, & Bunting, 2014). Additionally, complex WM measurements are better predictors of comprehension—average weighted effect size (*r*) between 0.41 to 0.52—than the simple WM tasks (Daneman & Merikle, 1996). Moreover, it is recommended that the WM task and its outcome measurement be matched in terms of modality (Grey et al., 2015)—for example if the outcome measurement is L2 listening, the WM task should be conducted aurally.

In looking at WM and the processing of Japanese case marking particles, Grey et al.'s study (2015) on *Japlish* was closely related to this study. Although the study matched the listening modality of the WM task and its outcome measurements, it used the simple WM measurement of L1 and L2 non-word repetition tasks finding no significant relationship between phonological working memory and the learning of *Japlish* case markings and word order. In order to have a better measurement on the WM task, this paper's study will instead use the complex WM task of the Listening Span Test, while matching the modality of the listening comprehension. The following section examines the relationship between the Listening Span Test and specific L2 listening outcomes.

2.2.2 The Role of Working Memory (WM) in L2 Listening Comprehension

In empirical studies that examine the relationship between L2 learning and WM tasks which utilize the Listening Span Test as a measurement of WM, participants with higher WM tended to perform better than those with lower WM in terms of modifying their output after feedback, learning new vocabulary and grammar patterns, and L2 listening comprehension scores. Mackey, Adams, Stafford, and Winke (2010) examined the relationship between learners' WM capacity and their production of modified output when they were given interaction feedback during conversation for L1 English L2 Spanish learners. Learners who had higher WM test scores tended to modify their output once receiving feedback. A similar study looked at the L2 development of question formatting in L1 Spanish L2 English older adult learners aged 65 to 89 years old through conversation (Mackey & Sachs, 2012). Although participants who had higher Listening Span Test scores showed a relatively strong relationship with immediate L2 development, it should be noted that this was a small-scale study and that the highest Listening Span Test score was lower than the lowest score in Mackey et al.'s study (2010) of university student participants.

Other studies using the Listening Span Test explored the relationship between WM and the ability to learn new vocabulary and grammar patterns in an artificial language testing L1 English speakers (Martin & Ellis, 2012). There was a positive correlation between WM and learning vocabulary and grammar patterns. Moreover, in a study looking at L1 Japanese L2 English learners in a TOEIC listening test, participants with higher WM tended to score higher in two dimensions of L2 listening comprehension—literal comprehension and inferential comprehension

(Miki, 2012). These studies suggest that participants with higher WM are more likely to be able to listen effectively by retaining more information, revising prior information, and accessing their long-term memory while listening, tasks which incur a greater demand in processing.

2.2.3 Foreign Language Aptitude: Sound Recognition Ability

Another variable that this paper's study examines is the relationship between sound recognition and the processing of scrambled sentences in the case marking language of Japanese. Sound recognition ability is one type of language aptitude. Language aptitude is a set of perceptual and cognitive abilities that predicts the rate learners acquire a foreign language in an effective and an efficient way (Saito, Suzukida, & Sun, 2018a). One of the most commonly known foreign language aptitude tests is the Modern Language Aptitude Test (MLAT) developed by Carroll and Sapon (1959). It has a high predictive validity, high content, and face validity and is used in many countries (Ortega, 2009). However, because MLAT scores are strongly related to how L2 learners acquire language in an early stage and in a foreign language classroom, there has been redevelopment and validation of new foreign language aptitude tests (Saito, 2017).

One of the widely cited computer-based aptitude tests is the LLAMA test developed by Meara (2005). The use of this test is supported by studies showing that L2 learners' lexicogrammar development is positively correlated with the LLAMA scores (Saito et al., 2018a). The LLAMA tests are loosely based on MLAT; however, over the years, they have diverged from the original on which they were based (Meara, 2005). The LLAMA tests are gender and language neutral (Rogers et al.,

2016) suggesting that they could be used in subjects with any language background. LLAMA-D, one of the subtests, utilizes verbal materials which are adapted from British-Columbian indigenous languages, rather than digits and symbols which are unrelated to natural language. Also, one component of LLAMA-D, sound recognition ability, is a new task which was not included in the MLAT (Granena, 2013a; Rogers et al., 2016). This component measures L2 learners' ability to recognize patterns in spoken language and is believed to measure aptitude for implicit language learning because there is no time during the task to work out the language rules or relationships (Granena, 2013b). In addition, not only does sound recognition ability help L2 learners acquire vocabulary but it also helps them to recognize small variations in the endings that mark grammatical features (Rogers et al., 2016) such as gender and/or number agreement (Granena, 2013b) as well as case marking particles. Since noticing case marking particles is the focus of this paper's study, LLAMA-D was chosen as a task to measure sound recognition ability.

2.2.4 The Role of Sound Recognition Ability in L2 Learning

In empirical studies that examined sound recognition ability in L2 learning measured by the LLAMA-D test, participants who had higher LLAMA-D score tended to perform well in certain aspects of their L2. For beginner L1 Spanish L2 English learners, those exhibiting higher levels of sound recognition ability tended to learn faster—LLAMA-D explaining 16 % of the variance (Artieda & Muñoz, 2016). Also L1 Japanese L2 English learners who had higher LLAMA-D scores had more substantial gains in developing their fluency and prosody in speaking their L2 throughout the academic year (Saito et al., 2018a) and had a better performance on the English /ɹ/ pronunciation (Saito, 2019). In addition, there was a significant relationship between LLAMA-D scores and recognizing agreement structures—such as gender, number, and subject-verb agreement—in L1 Chinese L2 Spanish learners who started learning at a younger age (Granena, 2013b). Even for late learners in L1 Swedish L2 French, there was a positive correlation between producing collocation and the LLAMA-D scores suggesting that there may be an association between sound recognition and the learning of vocabulary (Forsberg Lundell & Sandgren, 2013). Moreover, there was a slight correlation between knowledge of lexis and LLAMA-D scores for L1 Chinese L2 Spanish late learners as well (Granena & Long, 2013). Although the studies were on different L1 and L2 languages, they suggest that sound recognition ability measured by LLAMA-D seems to play a positive role in L2 learning.

2.3 Summary of the Review of the Literature

Case marking particles are an important aspect of Japanese grammar. Because of its flexible word order, for people to fully understand Japanese sentences, they would need to rely on the case marking particles to determine whether a noun phrase is a subject or an object in a sentence. In previous studies, there have been conflicting findings about whether scrambled sentences will incur processing costs for both NS Japanese and L2 Japanese learners. Much of the current literature that exists on L2 Japanese learners utilizes various types of measurements and outcomes, and with many focused on reading comprehension. In order to further evaluate whether L2 Japanese learners can comprehend case marking particles, canonical and scrambled sentences were used for this paper's study. In addition, a listening task was chosen given that it better assesses the real-world scenario in which L2 Japanese learners are most likely to encounter scrambled sentences.

Moreover, there are a limited number of studies on why some L2 Japanese learners have difficulty processing scrambled sentences. Individual differences play an important role in learning a foreign language and this may be related to comprehending case marking particles. For this paper's study, WM and sound recognition ability were chosen as individual differences given that learners with a greater WM are more likely to be able to retain information and access the previous knowledge effectively while listening, and this may be an essential factor in processing case marking particles correctly. Additionally, learners who possess higher sound recognition ability are more likely to successfully hear the various endings of grammatical features, such as case marking particles that occur at the end of noun phrases. These individual differences may make L2 Japanese learners more likely to be able to comprehend scrambled sentences accurately. Therefore, the present study will investigate how individual differences may have an effect on L2 Japanese learners' processing of case marking particles using canonical and scrambled sentence structure in a listening task.

Chapter 3: Research Questions and Hypotheses

Since there are a limited number of studies that compare the role of individual differences in processing of case marking particles for L2 Japanese learners, this study will examine the relationship between individual differences, especially in WM and sound recognition ability, and L2 Japanese learners' comprehension of scrambled Japanese sentences in a listening task. Based on the literature review, the following research questions and associated hypotheses and rationales were proposed.

Research Question 1: How does L2 Japanese learners' accuracy rate compare when processing canonical versus scrambled sentences in a listening task?

Hypothesis 1: Participants' accuracy rate will be lower for scrambled sentences compared with canonical sentences.

Although there is mixed evidence about the reaction times in both NS and L2 Japanese, L2 learners tended to have difficulty in sentence processing even in reading sentences—showing no subject-object asymmetry and not optimally utilizing case marking cues to read relative clause sentences (Mitsugi, MacWhinney, & Shirai, 2005). In addition, even when L1 English L2 Japanese learners demonstrated a good offline knowledge of case marking particles, they were unable to utilize them online in order to predict upcoming information in a visual world eye-tracking experiment (Mitsugi & MacWhinney, 2016). Moreover, although the studies were on reading comprehension, the accuracy rate for L2 Japanese learners became lower when scrambled sentences were read (Iwasaki, 2003; Shigenaga, 2012). This suggests that as a structure becomes more complicated, such as in scrambled sentences, L2 learners will incur longer processing times and thus may make more errors in scrambled sentences when under pressure, such as in a listening task.

Research Question 2: How does L2 Japanese learners' WM affect comprehension of scrambled Japanese sentences?

Hypothesis 2: Participants with larger WM capacity are more likely to be able to comprehend scrambled sentences than participants with smaller working memory capacity.

Since WM capacity influences the efficiency and the quality of real-time L2 processing (Miyake & Friedman, 1998), participants with larger WM will be more likely to be able to process scrambled sentences accurately. In empirical studies that examine the relationship between L2 learning and a WM task measured via the Listening Span Test, participants with higher WM tended to perform better than those with lower WM in terms of learning new vocabulary and grammar patterns (Mackey & Sachs, 2012; Martin & Ellis, 2012) and L2 listening comprehension scores (Miki, 2012). Therefore, L2 Japanese learners with higher WM will be more likely to be able to listen effectively by retaining case marking particles and retrieving previously learned grammatical information translating into better processing of Japanese scrambled sentences.

Research Question 3: How does L2 Japanese learners' sound recognition ability influence comprehension of scrambled Japanese sentences?

Hypothesis 3: Participants with greater sound recognition ability are more likely be able to comprehend scrambled sentences than participants with lower sound recognition ability.

Previous studies that examined sound recognition ability in L2 learning suggest that L2 learners with superior sound recognition ability tended to perform better in various aspects, such as vocabulary learning (Forsberg Lundell & Sandgren, 2013) and overall speed in learning the L2 (Artieda & Muñoz, 2016). In addition, sound recognition ability helps L2 learners recognize small variations in the endings that marks grammatical features (Rogers et al., 2016) such as gender and/or number agreement (Granena, 2013b) and case marking particles. Since the focus of this study is on comprehending case marking particles properly, if L2 Japanese learners have a greater sound recognition ability, they will be more likely to differentiate the case marking particles which occur at the end of noun phrases. In addition, they will be more likely to be utilize these particles to effectively comprehend scrambled sentences.

The research questions and the hypotheses proposed are based on the existing literature extrapolated to L2 Japanese learners' ability to comprehend scrambled sentences in a listening task. In order to better examine the role of individual differences, especially WM and sound recognition ability, on processing case marking particles, the present study was conducted.

Chapter 4: Methods

4.1 Participants

Seventy-one Japanese L2 learners participated in the study (41 females, 29 males, and 1 other). All participants reported having good hearing and normal or corrected-to-normal vision. They were undergraduate and graduate students in the Washington D.C., Maryland, and Virginia area who are or were learning Japanese as a foreign language. Participants were enrolled in a 3rd year Japanese course level at college (n = 40) or above (i.e. 4th year or completed all course works offered in Japanese; n = 31) and many had experience studying abroad in Japan (n = 54). The mean age at the time of the study was 21.63 years old (SD = 2.99, Min = 18, and Max = 37).

Before the experiment, all participants completed a background information questionnaire. 39 participants spoke only English at home, while 21 participants spoke English plus another language at home, such as Bengali, Cebuano, Chinese, French, Japanese, Korean, Russian, Slovak, Spanish, Tagalog, and Telugu. 11 participants exclusively spoke a non-English language at home, such as Chinese, Thai, and Vietnamese. The mean age participants started learning Japanese was 16.35 (SD = 4.45), and the mean duration participants had been learning Japanese in college was 2.8 years (SD = 1.12).

Participants were also asked to self-rate their proficiency level in Japanese and, if English was not their home language, their proficiency level in English. Proficiency levels were rated on a scale from 1 to 5 (1 = with difficulty; 5 = easily) in each of the following components: speaking, listening, writing, and reading skills.

Table 1 shows participants' mean self-rated level of proficiency in Japanese. Among the 11 participants who exclusively spoke a non-English language at home, their age of acquisition of English was 6.27 (*SD* = 1.35), and the majority chose 4 or 5 for all the criteria in their self-rated English proficiency level which had the same prompt as their self-rated Japanese proficiency level.

Table 1. *Self-Ratings of Japanese Language Skills*

	Mean	***SD***
Converse with friends	3.53	1.16
Converse with professors	3.25	1.04
Discuss social issues	2.55	1.07
Watch TV without subtitles	2.97	1.09
Read short blogs	3.14	1.09
Read newspaper articles	2.34	1.06
Write emails to a professor	3.56	0.92
Write a short essay	3.29	0.95

Note—*N* = 71. Scale: 1 = *with difficulty*; 5 = *easily*.

4.2 Materials

In the following section, the materials used in this study are explained in detail with regard to the task's focus, how the task was conducted, and what the reliability of the task was.

4.2.1 Target Task

All the target sentences for this study were grammatically correct canonical and scrambled sentences using the nominative case marker -*ga* and accusative case marker -*o*. Sentences like the following were constructed. The full set of sentences can be found in Appendix A.

Canonical Sentence Example:

3) John-ga Mary-o suupaa-de mimashita.

-NOM -ACC supermarket-LOC see-PAST

'John saw Mary at the supermarket.'

Scrambled Sentence Example:

4) Mary-o John-ga suupaa-de mimashita.

-ACC -NOM supermarket-LOC see-PAST

'John saw Mary at the supermarket.'

In the target sentences, to make both the nominative and the accusative cases be animate nouns, only Japanese and English proper names were used. This was to minimize the likelihood that participants would utilize methods other than the case marking particles as cues to comprehend the sentences. If a sentence were to have both an animate and an inanimate noun, participants would be more likely to guess the meaning from the context rather than from the case marking particles. For example, in order to make Sentence # 3 into a scrambled sentence, 'John-*ga*' and 'Mary-*o*' was switched, making it into Sentence # 4. Although the locative case marker -*de* could be scrambled as well, as the purpose of this study was to see whether participants can use case marking particles to comprehend sentences, only the noun phrases that had -*ga* and -*o* particles were switched to make scrambled sentences. Other than the locative case marker -*de*, the dative case marker -*ni* was used to make the sentence longer and to decrease the probability of participants easily identifying the task.

A total of 64 sentences were constructed—32 target sentences and 32 filler sentences. The 32 target sentences were split into two equivalent sub-lists—List A and B—for the purpose of using DMDX. At this point, scrambled sentences were created by switching the order of the *-ga* and *-o* marked noun phrases. To ensure that participants in the experiment would not listen to two sentences with the same combination of noun phrases and verbs, the scrambled sentences which were created from the canonical sentences on List A would appear exclusively on List B and vice versa. Thus, two lists were created which consisted of 16 sentences each of both sentence types.

For the filler sentences, half were in SOV sentence structures such as in example 5. One thing to note is that the *-o* particle was used as well after inanimate nouns in the filler sentences such as the following sentence because it was thought it would not affect participants' comprehension even if they were not able to utilize case marking particles.

5) Mary-ga hambaagaa-o resutoran-de tabemashita.

-NOM hamburger-ACC restaurant-LOC eat-PAST

'Mary ate a hamburger at the restaurant.'

The other half of the filler sentences were in different sentence structures so that participants would not focus on the sentence structure but rather on the meaning of the sentence. Whereas the first half of the filler sentences utilized the case marker *-o* with inanimate nouns, the second half utilized a variety of sentence structures employing varying case marking particles such as with the topicalizer *-wa*, dative case marker *-ni*, duration of origin or motion case marker *-kara*, and comparison particles *-yori*. There were also sentences with adjectives in order to minimize participants exclusively focusing on the case marking particles. These filler sentences further substantiate participants' Japanese proficiency in conjunction with the SPOT Test Version B (see below for more detail about the SPOT Test Version B).

All the sentences were followed by a yes/no comprehension question in English (i.e. "Did John see Mary at the supermarket?"). These comprehension questions were formed in such a way that the first noun heard in the target task was the first noun in the comprehension question as in the examples below.

Canonical Sentence Example:

6) Sentence heard: John-ga Mary-o suupaa-de mimashita.

-NOM -ACC supermarket-LOC see-PAST

'John saw Mary at the supermarket.'

Comprehension question: Did John see Mary at the supermarket? / Answer: Yes

Scrambled Sentence Example:

7) Sentence heard: Mary-o John-ga suupaa-de mimashita.

-ACC -NOM supermarket-LOC see-PAST

'John saw Mary at the supermarket.'

Comprehension question: Did Mary see John at the supermarket? / Answer: No

Since the comprehension question was made to be a yes/no question, in order to account for the fact that some participants may have guessed, a confidence level question—guess, somewhat confident, absolutely confident—was conducted as well. The sentences were recorded by a female native Japanese speaker and the task was conducted using DMDX (Forster & Forster, 2003).

Lastly, because there were two lists, participants were randomly assigned to either List A or B. According to the Rasch Analysis, the person reliability in List A was 0.78, while in List B, it was 0.81; and the item reliability was 0.87 and 0.83 respectively. Both the person and the item reliability in both lists are similar indicating that there was no difference in the difficulty among the two lists.

4.2.2 SPOT Test Version B

To ensure that participants were of the intermediate-level, SPOT Test Version B (paper based) was used to measure their Japanese proficiency. This test was developed by the International Student Center at the University of Tsukuba in Japan[2] and tests participants' listening skill as well as grammatical knowledge (Kobayashi, 2016). The test consisted of 60 Japanese sentences with one missing morpheme which is left blank. Participants listened the Japanese sentences and had to fill in the blank with the appropriate Hiragana characters. The reliability of the task based on Kuder-Richardson Formula 21 (KR-21) was 0.88.

4.2.3 Listening Span Task

Since this study is on participants' listening skills, auditory working memory ability was measured through the Listening-Span Task developed by Mackey, Adams, Stafford, & Winke (2010) and the materials were downloaded from IRIS[3]. The task consists of 48 sentences read aloud in sets of three, four, or five sentences. While participants were listening, they needed to decide whether the sentence made sense or was nonsense and if it was grammatical or ungrammatical. Once participants were finished listening to a sentence set, they were asked to write down the last word in

[2] For more information, please go to their website (http://ttbj-tsukuba.org/index.html).
[3] For the Listening Span Task, please visit the IRIS website (https://www.iris-database.org/iris/app/home/detail?id=york%3a806454&ref=search).

each sentence. The number of sentences increased throughout this task up to 5 sentences in a set with a total of 12 sets.

Working memory was determined with aggregated scoring on both parts—participants' determination of whether the sentence both made sense and was grammatical as well as recalling the sentences' final words. The scoring procedure was identical as in Mackey et al. (2010)—half-point for correct judgement of sense or nonsense, half-point for correct judgement of grammatical or ungrammatical, and one point for correct recalled words. For the recalled words, if participants added or deleted a plural morpheme, that was counted as one point as well. There were 48 sentences on the test corresponding to 48 points possible for processing (sense/nonsense and grammatical/ungrammatical) and 48 points possible for recalling the last words for a total of 96 points possible. The KR-21 reliability was 0.47 for processing, 0.83 for recall, and 0.80 for the total score. Since the reliability for the processing section was low, it was not included in the final analysis.

4.2.4 LLAMA-D

To measure participants' sound recognition ability, LLAMA-D test was used. This was developed by Meara (2005) and was downloaded from its website[4]. The task consisted of two parts. The first part was the input in which participants were asked to listen to a string of 10 foreign words. They were instructed to try to remember those 10 words. Once participants finished listening, they heard another word and had to decide whether that word was in the initial word string. This was

[4] For the LLAMA tests, please visit the website (http://www.lognostics.co.uk/tools/llama/).

used to test whether participants are sensitive to recognizing case marking particles when listening.

Since the program for LLAMA-D does not automatically record the individual item responses, the beep sounds indicating the correct or the incorrect responses were audio recorded using the free software Audacity[5]. Each test item was then scored correct or incorrect with a maximum score of 30. The person reliability based on the Rasch Analysis was subsequently determined to be 0.09. Due to the very low reliability among the participants, LLAMA-D was excluded from the final analysis.

4.3 Procedures

All participants were tested individually, and the experimenter remained in the same room for the duration of the experiment. First, participants were asked to read and sign a consent form in which an overview of the study was presented. The study was then introduced to them as a project which was investigating participants' listening skill and reaction time. After the consent form was read and signed, a background questionnaire was administered which inquired about subjects' age, gender, Japanese course level and/or Japanese Language Proficiency Test (JLPT) level, language(s) spoken at home, other language(s) studied, length of study of Japanese including study abroad, and the self-rating of their own Japanese language level, and their English proficiency for non-native English speakers. For the complete background questionnaire, see Appendix B.

[5] For the Audacity software, please visit the website (https://www.audacityteam.org/).

Once the participants had completed the consent form and background questionnaire, they were guided to the following 4 tasks: Target Task, SPOT Test Version B, Listening Span Task, and LLAMA-D. For all four tasks, participants listened to audio recordings on a computer through headphones. The contents of each task and how it was administered are described as follows.

In the Target Task, participants sat in front of a laptop computer and were first given instructions both visually and orally on how to do the task. Each trial began with a pretrial warning, an asterisk (*) that appeared for 500 milliseconds (ms) in the middle of the screen, and was followed by the audio recording of the first Japanese sentence. After hearing a Japanese sentence, an English comprehension question appeared in the middle of the computer screen. Participants were instructed to press the right-shift key that was labeled 'YES' and the left-shift key that was labeled 'NO' in order to answer the comprehension question. The 'YES' and 'NO' labels were also provided on the bottom corners of the screen so that the participants could have their fingers on the keys but still know which one correspond to 'YES' or 'NO.' They had 10,000 ms to answer the comprehension questions. After the comprehension question, they were further asked to indicate the level of confidence they had in answering the comprehension question. They used the number keys 1-3 on a keyboard, which corresponded to 1 = 'guessed', 2 = 'somewhat confident', and 3 = 'very confident.' This code was projected on the screen after each question for the confidence rating part so that participants did not have to memorize the numbering system in order to minimize their memory load and be able to focus on the main task, the English comprehension question. Participants had 15,000 ms to respond to their level of confidence. They were instructed to respond as quickly and as accurately as possible. They were also told that if they do not press any buttons, the questions would automatically move on after the allotted time. There were 4 practice items followed by 64 items—16 canonical, 16 scrambled, and 32 filler sentences. The order of the presentation was randomized, and the participants were able to listen to each sentence once. Figure 1 summarizes the entire procedure of the task using DMDX (Forster & Forster, 2003).

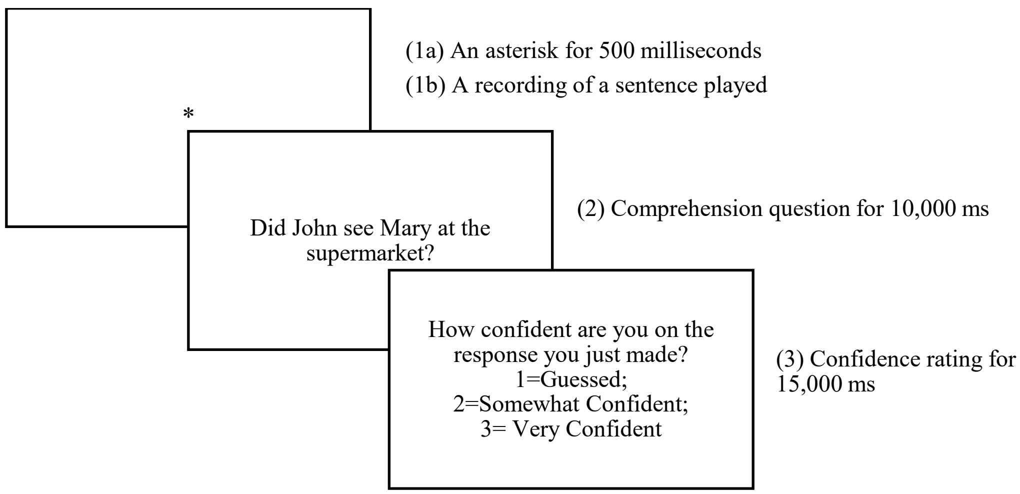

Figure 1. Procedure for DMDX

After the Target Task, the laptop computer was removed from the participants and they were given a worksheet and a pen to complete the SPOT Test Version B. Participants were instructed to listen to the audio and to fill in the blank with the appropriate Hiragana character. They had 10 practice items, then 60 items total.

For the Listening Span Task, since the instructions were pre-recorded, participants were told to listen to them individually. If they had any questions about the instructions, they were told to ask for clarification. If they did, the audio was paused and was explained in detail. Participants had one set of sentences as a practice set followed by 12 sets of sentences. Participants were given a worksheet and a pen to complete this section as well.

In the LLAMA-D section, participants were provided a laptop computer again and were instructed both visually and orally on how to use the program on the computer. There was no practice phase in this section. Once participants finished listening to the initial 10 foreign words, there were 30 additional words to which participants listened one at a time. After listening to each word, they responded whether that word was in the initial word-string or not by clicking on the appropriate response on the computer screen using a mouse.

At the end of the experiment, participants were asked to complete an Exit Survey asking what they were focusing on when doing the Target Task and the SPOT Test Version B. They were also asked their methods on improving their listening skills and what they know about Japanese sentences structure. For the complete exit survey, see Appendix C. Upon completing the Exit Survey, participants were thanked for their participation and received US$15 as compensation. The entire process took approximately 60 minutes to complete and participants were allowed to take breaks as desired throughout the experiment. The overall procedures are shown in Table 2.

4.4 Analysis

The primary analysis for this study focused on the relationship between individual differences and the accuracy of scrambled sentences. For this part, Generalized Linear Mixed Effects Models analyses were conducted using a software

Table 2. *Procedure of the Study*

Task	Time (Minutes)
Consent Form and Background Questionnaire	7
Target Task	18
SPOT Test Version B	7
Listening Span Task	18
LLAMA-D	5
Exit Survey and Compensation	5
	Total = 60

Supermix. The dependent variable is the accuracy and the independent variable is the sentence type—canonical, scrambled, or filler. The filler sentences were included in the analysis as sentences with meanings that participants would be able to understand through context even if they were not able to utilize case marking particles fully. Thus, this sentence type served as another variable to compare with scrambled and canonical sentences. Additional independent variables are the participants' working memory measured through the Listening-Span Task and Japanese proficiency level measured by the SPOT Test Version B. For the Listening-Span Task, the total score as well as the score for recalling the last word of each sentence were used. In addition, to account for the differences in the level of difficulty of item types and proficiency level among participants, items were indicated as Level 2 ID and the participants as Level 3 ID. The analysis will focus on what combination of factors—proficiency, WM, or proficiency x WM—influence the probability of understanding scrambled sentences.

The secondary analysis focused on evaluating the accuracy between comprehending canonical and scrambled sentences. For this section, within-subject paired *t*-test using SPSS was used to measure the difference in accuracy for comprehending canonical versus scrambled sentences. In addition, reaction times between these two types of sentences were also analyzed via within-subject paired *t*-test using SPSS to determine whether there was any significant slowdown when listening to scrambled sentences. The differences were judged to be significant for *p* value of less than 0.05.

Chapter 5: Results

5.1 Accuracy of Scrambled Sentences

The results of participants' mean accuracy and reaction times (RTs) are summarized in Table 3. The scrambled sentences had lower participant accuracy and a longer participant RT compared with canonical and filler sentences. The distribution of the total accuracy rate of all the participants on each sentence type are indicated in Figure 2.

Table 3. *Mean Accuracy and Accuracy Rate (in parenthesis) and Reaction Times (RTs in Milliseconds) for Each Sentence Type*

	Accuracy		RT	
	Mean Difference	*SD*	Mean Difference	*SD*
Canonical (n = 16)	14.34 (89.61 %)	1.69	3433.61	1070.52
Scrambled (n = 16)	5.93 (37.06 %)	4.46	3710.53	1075.65
Filler (n = 32)	29.44 (91.99 %)	2.35	2879.73	851.97

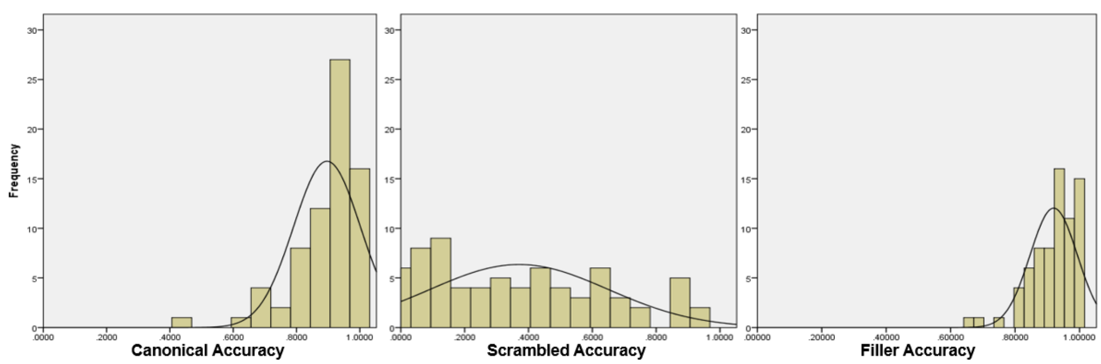

Figure 2. Histogram of Participants' Accuracy Rate among Sentence Types

As can be seen from Figure 2, participants' accuracy rate for scrambled sentences is skewed to the lower end—most participants scored on the lower end and very few scored high. On the other hand, the canonical and filler sentences are towards the higher end of accuracy rate with the majority of participants scoring close to 100%. To examine whether there was a significant difference in the accuracy rate among the sentence types, a within-subject paired *t*-test was conducted using SPSS. Between the canonical and scrambled sentences, t (70)=14.47, $p < 0.001$ suggesting that there was a significant difference in the accuracy rate. Between the filler and scrambled sentences, t (70)=17.33, $p < 0.001$ also suggests that there was a significant difference between the accuracy rate among these two sentence types. Lastly, between the filler and canonical sentences, it showed that t (70)=1.77, $p = 0.082$, suggesting that the difference in participant accuracy for these two sentence types was not statistically significant. The overall results are reported in Table 4 and indicate that scrambled sentences are harder for L2 Japanese learners to comprehend than canonical and filler sentences.

Table 4. *Paired Sample t-Test Among Sentence Types on Accuracy Rate*

	Mean	*SD*	*t*	df	Sig. (2-tailed)
Canonical - Scrambled	0.53	0.31	14.47	70	0.000*
Filler - Scrambled	0.55	0.27	17.33	70	0.000*
Filler – Canonical	0.02	0.11	-1.77	70	0.082

Note. * indicates $p < 0.001$.

5.2 Individual Differences

The primary outcome of interest in this study was to determine whether individual differences affect L2 Japanese learners' comprehension of Japanese scrambled sentences. Using the Generalized Linear Mixed Effects Model analysis on the Supermix software, the analysis focused on what combination of factors—proficiency, WM, or proficiency x WM—affect learners' understanding scrambled sentences.

5.2.1 Proficiency: SPOT Test

Participants' proficiency was measured by the SPOT Test Version B (henceforth termed SPOT). The mean score of the test was 55.08 (SD = 5.71, Min = 30, and Max = 60). The maximum score possible was 60. Figure 3 shows the distribution of the participants' SPOT score indicating that many people scored on the high end of the test. This suggests a high average Japanese listening ability among the participants.

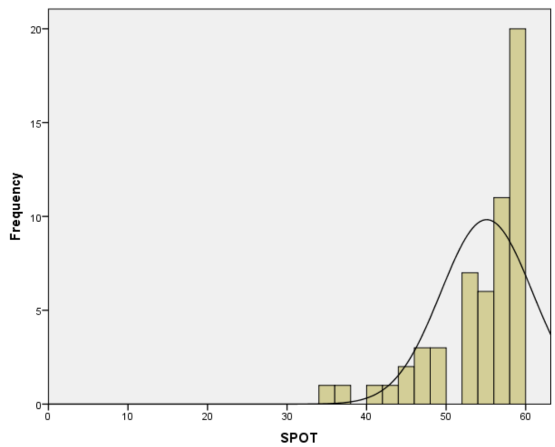

Figure 3. Histogram of the Participants' Accuracy for SPOT Test Version B

The Generalized Linear Mixed Effects Model for SPOT and scrambled sentences showed that the p values were $p < 0.001$ for both SPOT and scrambled sentences; while the odds ratios were 1.06 and 0.05 respectively. From the odds ratios, the probability of being accurate was calculated using the following formula:

$$\text{Probability of Being Accurate} = \frac{Odds\ Ratio}{1\ + Odds\ Ratio}$$

The probability of being accurate was 0.51 for SPOT and 0.05 for scrambled sentences. The score on the probability of being accurate can be interpreted as if there were 100 SPOT Test questions, the probability that the participants will answer them correctly is 0.51 or 51 %, while the probability for answering scrambled sentences correctly would be 0.05 or 5 %. This indicate that the probability of accurately comprehending scrambled sentences increases 5 % for every correctly scored SPOT test. This suggests that even as participants' proficiency in Japanese becomes higher, their likelihood of comprehending scrambled sentences only minimally increases demonstrative of the difficulty with which participants' process scrambled sentences. For SPOT and canonical sentences, the p values were $p < 0.001$ for SPOT and canonical sentences while the odds ratios were 1.04 and 3.27 respectively. The probability of being correct was 0.51 for SPOT and 0.77 for canonical sentences. Similar to interpreting the probability of being accurate for scrambled sentences, if there were 100 canonical sentences, the probability that the participants will answer them correctly would be 0.77 or 77 %. This indicates that the probability of accurately comprehending canonical sentences increases 77 % for each additional correct response on the SPOT test suggesting that as participants' proficiency becomes higher, it becomes easier to comprehend canonical sentences. For SPOT and filler sentences, the p values were $p < 0.001$ for both SPOT and filler sentences; while the odds ratios were 1.05 and 6.89 respectively. The probability of being correct was 0.51 for SPOT and 0.87 for filler sentences. Since there is 87 % probability of being accurate for each additional correct response on the SPOT test, it indicates that as the participants' proficiency becomes higher, it becomes relatively easier to comprehend filler sentences. Figure 4 describes the overview of the probability of being accurate by sentence type. Overall, proficiency did not play a large role in the comprehension of scrambled sentences suggesting that even if participants have high proficiency in

Japanese, they may still have a hard time processing case marking particles effectively in a listening task.

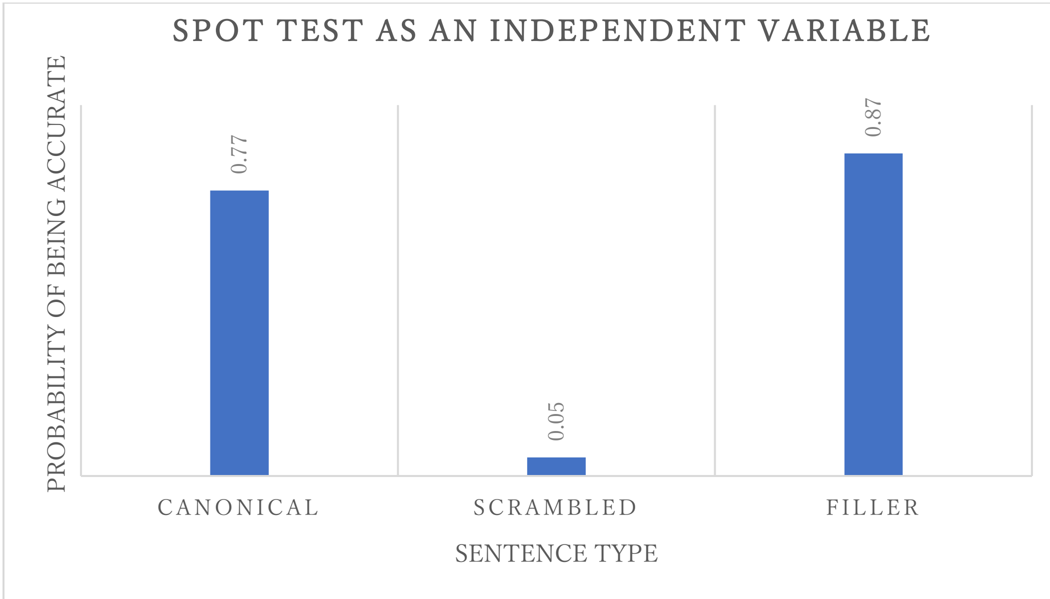

Figure 4. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' SPOT Test Version B Scores

5.2.2 Working Memory

Since proficiency did not play a large role in identifying case marking when the sentences were scrambled, WM was used to analyze how it may mediate accuracy and the recognition of different types of sentence structures. As mentioned earlier, two types of WM were used: recalling the last words of the sentences and the total scores from the Listening-Span Task. The results are described below.

5.2.3 WM Recall

The score on the subsection of the Listening-Span Test in which participants were required to recall the last words in each sentence was used for this part of the analysis (hence WM Recall). The mean score on the WM Recall was 34.17 (*SD* = 7.54, Min = 11.5, and Max = 48). The maximum score possible was 48 points, and as can be seen from Figure 5, there is a normal distribution in the higher end but, at the same time, there are slight variations in their WM Recall scores suggesting that there exists a difference among participants.

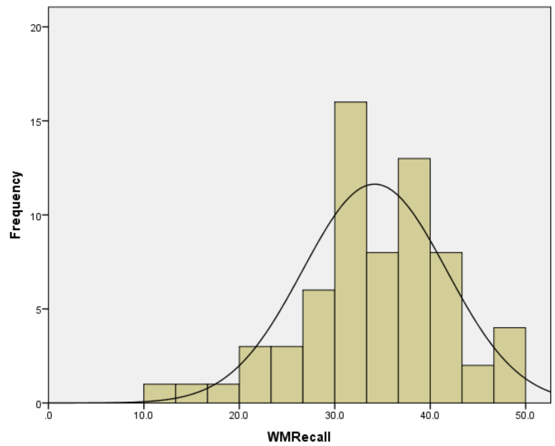

Figure 5. Histogram of Participants' Accuracy for WM Recall

For the Generalized Linear Mixed Effects Model for WM Recall and the scrambled sentences, the *p* values were $p < 0.001$ for WM Recall and scrambled sentences, while the odds ratios were 1.05 and 0.05 respectively. The probability of being correct was 0.51 for WM Recall and 0.05 for scrambled sentences suggesting that even if participants' WM Recall scores become higher, scrambled sentences still remain difficult to comprehend. For WM Recall and canonical sentences, the *p* values were $p < 0.001$ for WM Recall and canonical sentences, while the odds ratios are 1.03 and 3.30 respectively. The probability of being correct was 0.51 for WM Recall and 0.77 for canonical sentences indicating that the canonical sentences become easier for participants as WM Recall becomes higher. For WM Recall and filler sentences, the *p* values were $p < 0.001$ for both WM Recall and filler sentences, while the odds ratios were 1.03 and 6.56 respectively. The probability of being correct was 0.51 for WM Recall and 0.87 for filler sentences suggesting that as the WM Recall scores become higher it becomes easier for participants to comprehend filler sentences. Figure 6 depicts of the probability of being accurate among sentence types and shows a similar pattern to the SPOT Test.

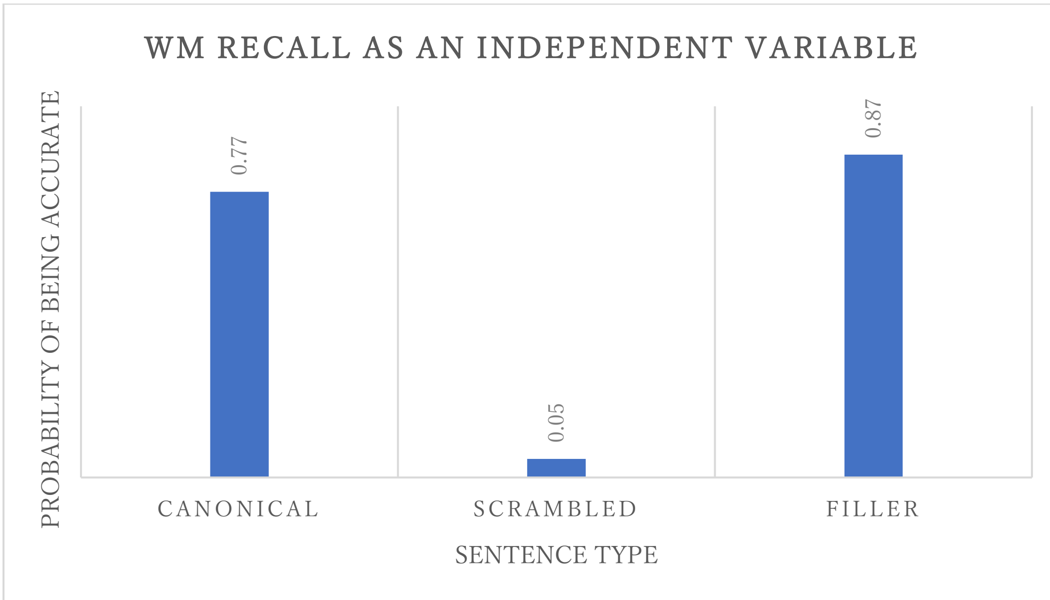

Figure 6. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Recall Scores

5.2.4 WM Total

The mean score of the total WM scores from the Listening-Span Task (henceforth WM Total) was 73.75 (*SD*=9.36, Min = 50, and Max =91). The maximum score possible was 96 points. Although Figure 7 shows that the distribution is slightly skewed to the higher end, there is a normal distribution among participants suggesting that there are some variations among them.

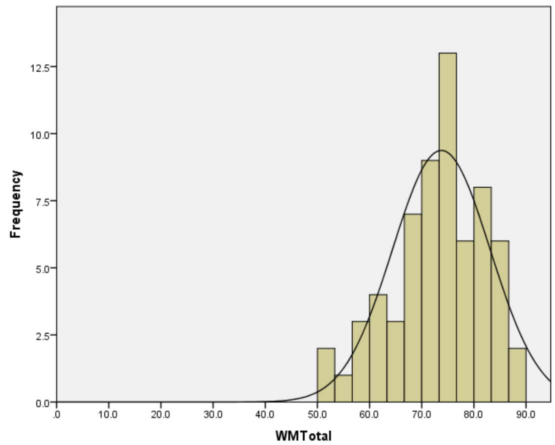

Figure 7. Histogram of Participants' Accuracy for WM Total

The Generalized Linear Mixed Effects Model for WM Total and the scrambling sentences showed that the *p* values were $p < 0.001$ for WM Total and scrambled sentences, while the odds ratios were 1.04 and 0.05 respectively. The probability of being correct was 0.51 for WM Total and 0.05 for scrambled sentences indicating that even if participants have higher WM Total scores, it was still challenging to process scrambled sentences. For WM Total and canonical sentences, the *p* values were $p < 0.001$ for WM Total and canonical sentences, while the odds ratios are 1.03 and 3.31 respectively. The probability of being correct was 0.51 for WM Total and 0.77 for canonical sentences indicating that canonical sentences were relatively easier for participants to process compared with scrambled sentences. For WM Total and filler sentences, the *p* values were $p < 0.001$ for WM Total and filler sentences, while the odds ratios were 1.04 and 6.56 respectively. The probability of being correct was 0.51 for WM Total and 0.87 for filler sentences indicating that as the WM Total becomes higher, it becomes relatively easier for participants to comprehend filler sentences. Figure 8 depicts the probability of being accurate among sentence types and shows a similar pattern as SPOT and WM Recall.

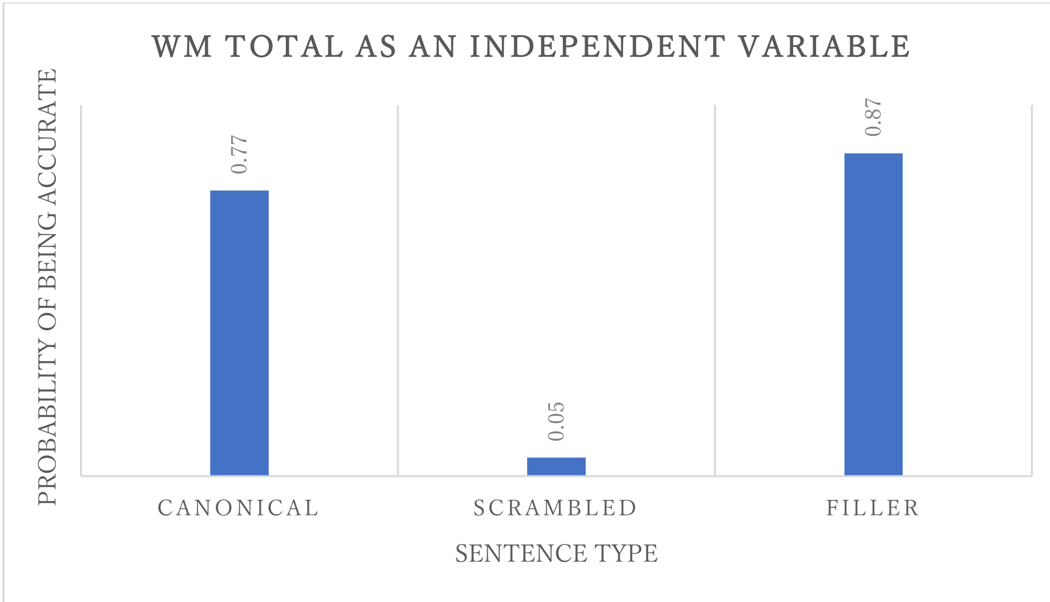

Figure 8. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Total Scores

Overall, each variable had a similar impact on the probability for being accurate for each sentence type. These results suggest that participants struggled to comprehend scrambled sentences even if they had high proficiency or WM.

5.2.5 WM Recall and SPOT

This section examines whether having both higher WM Recall scores and SPOT scores will affect participants' processing of scrambled sentences. A Pearson correlation was computed to assess the relationship between WM Recall and SPOT. There was a positive correlation between the two variables, r = 0.23, n = 66, *p* = 0.06. Participants who scored high on the WM Recall also tended to score high in the SPOT test, although this trend was not statistically significant. A scatterplot is shown in Figure 9.

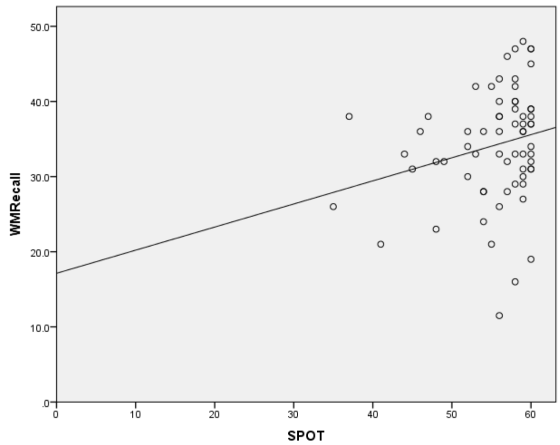

Figure 9. Scatterplot of Participants' Accuracy in SPOT and WM Recall

The Generalized Linear Mixed Effects Model for WM Recall, SPOT, and the scrambled sentences showed that the p values were $p < 0.001$ for all three variables, while the odds ratios were 1.04, 1.05, and 0.05 respectively. The probability of being correct was 0.51 for WM Recall and SPOT and 0.05 for scrambled sentences suggesting that even if participants' scores on both WM Recall and SPOT become higher, it is still difficult for participants to comprehend scrambled sentences. For WM Recall, SPOT, and canonical sentences, the p values were $p < 0.001$ for all three variables, while the odds ratios are 1.03, 1.03 and 3.33 respectively. The probability of being correct was 0.51 for WM Recall and SPOT and 0.77 for canonical sentences indicating that as WM Recall and SPOT scores become higher, canonical sentences become easier for participants to comprehend. For WM Recall, SPOT, and filler sentences, the p values were $p < 0.001$ for the three variables, while the odds ratios were 1.03, 1.04, and 6.66 respectively. The probability of being correct is 0.51 for WM Recall and SPOT and 0.87 for filler sentences indicating that as the score for WM Recall and SPOT becomes higher, it becomes relatively easier for participants to comprehend filler sentences. Figure 10 displays the probability of being accurate by sentence type. Even if participants had both higher WM Recall and SPOT scores, they still had difficulty processing scrambled sentences.

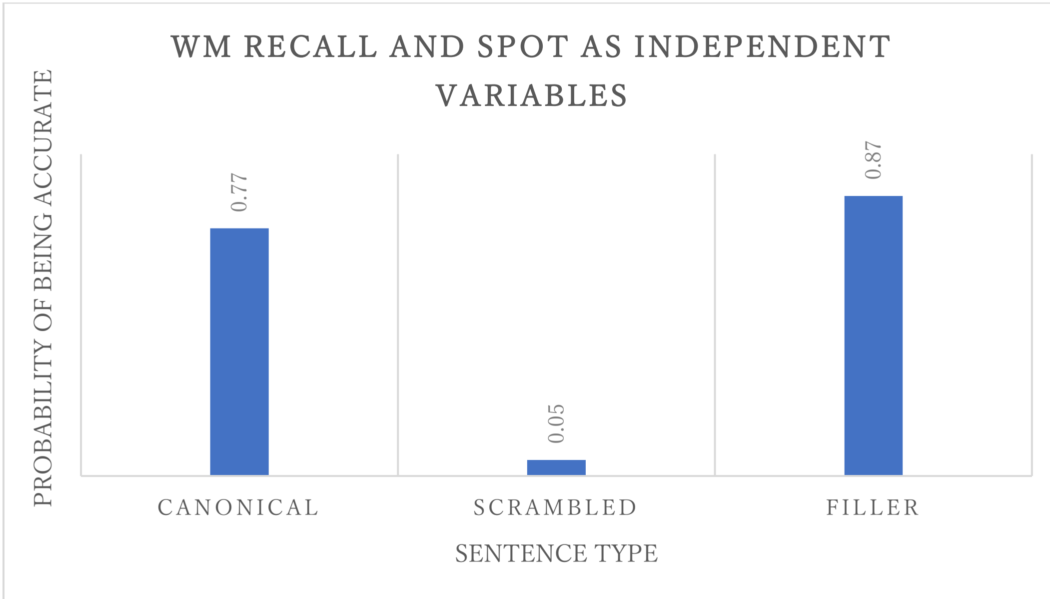

Figure 10. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Recall and SPOT Test Version B Scores

5.2.6 WM Total and SPOT

A similar analysis was conducted for WM Total and SPOT. A Pearson correlation indicated that there was a weakly positive but significant correlation between the two variables, r = 0.25, n = 66, p = 0.05. A scatterplot shown in Figure 11 shows that participants who scored high on the WM Total also tended to score high in the SPOT test, which is a similar patter with WM Recall and SPOT. Both WM Recall and WM Total had a positive correlation with SPOT test. This may be due to the fact that participants who have higher WM are better at retaining information, accessing previously learned materials effectively, and processing multiple information at once (Gathercole, 2006; Miyake & Friedman, 1998). They may therefore be more likely to perform better in their proficiency test as well.

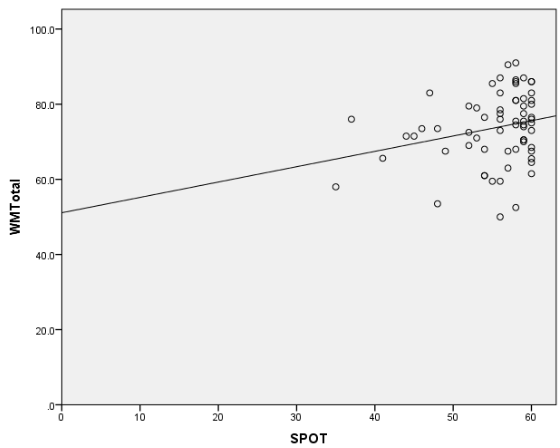

Figure 11. Scatterplot of Participants' Accuracy on WM Total and SPOT (r = 0.25)

In the Generalized Linear Mixed Effects Model, the WM Total, SPOT, and the scrambling sentences showed that the *p* values were $p < 0.001$ for all three variables, while the odds ratios were 1.03, 1.05, and 0.05 respectively. The probability of being accurate was 0.51 for WM Total and SPOT and 0.05 for scrambled sentences suggesting that even if participants' score on both WM Total and SPOT became higher, it was still difficult for participants to comprehend scrambled sentences. For WM Total, SPOT, and canonical sentences, the *p* values were $p < 0.001$ for all three variables, while the odds ratios were 1.02, 1.03, and 3.33 respectively. The probability of being correct was 0.50 for WM Total, 0.51 for SPOT, and 0.77 for canonical sentences indicating that canonical sentences were relatively easier for participants if they had high WM Total and SPOT scores. For WM Total, SPOT, and filler sentences, the *p* values were $p < 0.001$ for all three variables, while the odds ratios were 1.03, 1.04, and 6.67 respectively. The probability of being correct were 0.51 for WM Total and SPOT, and 0.87 for filler sentences indicating that the filler sentences were relatively easy for participants who scored high in both WM Total and SPOT. Figure 12 describes the overview of the probability of being accurate by sentence type. Even if participants had both higher WM Total and SPOT scores, they still had difficulty processing scrambled sentences.

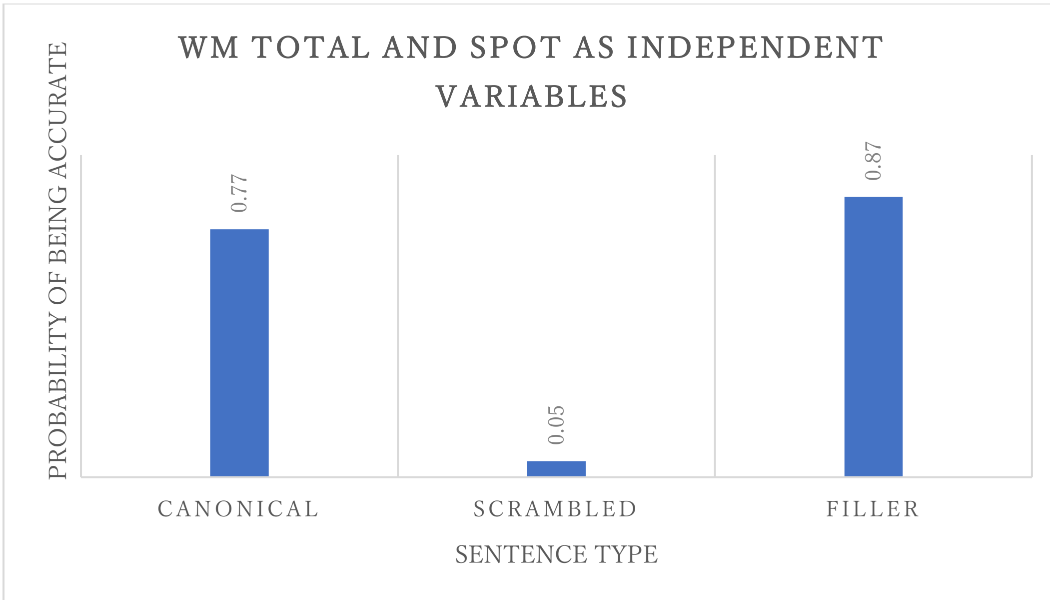

Figure 12. Bar Graph of the Probability of Being Accurate for Each Sentence Type Based on Participants' WM Total and SPOT Test Version B Scores

Given that the probability of being accurate remained almost constant even after combining WM Recall and SPOT, and WM Total and SPOT together, the data suggests that even in participants with high scores on WM Recall, WM Total, and SPOT, their comprehension of scrambled sentences becomes constrained compared with canonical and filler sentences. This suggests that participants cannot process case marking particles as efficiently regardless of proficiency and WM in a listening task.

5.2.7 Interaction Between Variables

To identify whether there was any interaction between sentence type and the type of individual differences, the interaction was analyzed as well. Table 5 shows the odds ratios, and the probability of being accurate for each interaction.

Table 5. *The Odds Ratios and Probability of Being Accurate (in parenthesis) of the Interaction among Sentence Type and SPOT, WM Recall, and WM Total*

	SPOT	WM Recall	WM Total
Canonical	1.02 (0.50) *	1.04 (0.51) *	1.02 (0.50) *
Scrambled	0.95 (0.49) *	0.93 (0.49) *	0.96 (0.49) *
Filler	1.04 (0.51) *	1.05 (0.51) *	1.03 (0.51) *

Note. * indicates the *p* value is $p < 0.001$.

Scrambled sentences unto themselves have a low probability of being accurate (0.05), while SPOT by itself has a probability of being accurate of 0.51. However, if there is an interaction between scrambled sentences and SPOT, the probability becomes slightly lower (0.49) compared with SPOT's probability of being accurate individually indicating that scrambled sentences constrain participants' proficiency and result in a slight decrease in their comprehension outcome. A similar pattern can be observed with WM Recall and WM Total. These results suggest that scrambled sentences will interact negatively with proficiency and WM. The interaction with scrambled sentences is indicated in Figure 13.

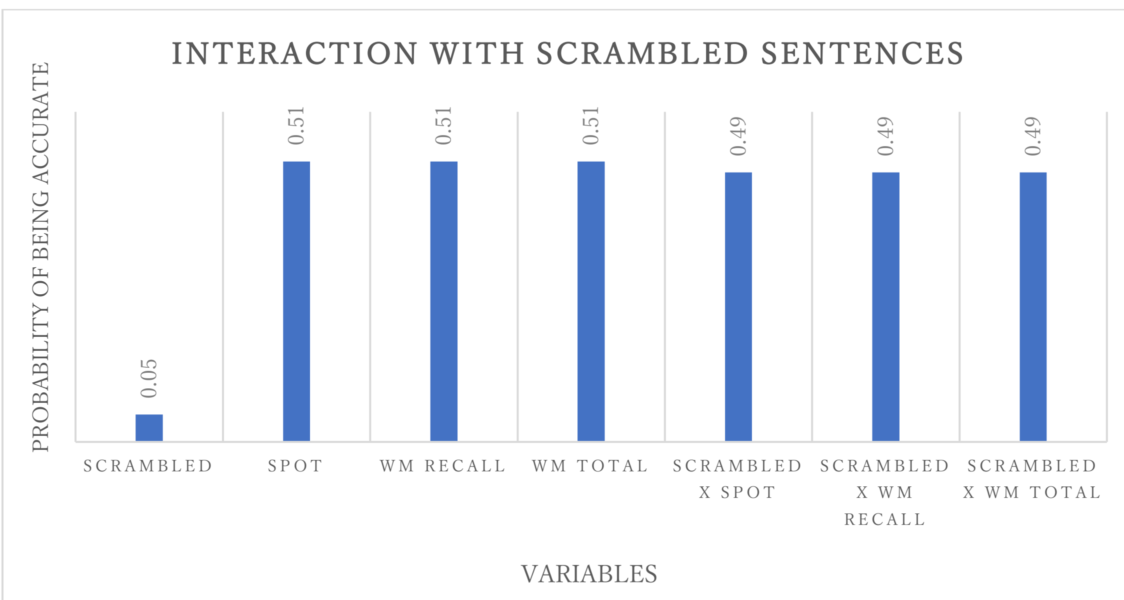

Figure 13. Bar Graph of the Probability of Being Accurate in Scrambled Sentences and Comparing its Interaction with SPOT, WM Recall, and WM Total

On the other hand, the odds ratios and the probability of being accurate remain the same when there are interactions between canonical sentences and proficiency, canonical sentences and WM, filler sentences and proficiency, or filler sentences and WM compared with when the variables were examined individually. This suggests SPOT or WM do not affect the comprehension of canonical sentences or filler sentences.

5.3 Reaction Time of Scrambled Sentences

The mean reaction times (RTs) were 3433.61 milliseconds (ms) for canonical sentences, 3710.53 ms for scrambled sentences, and 2879.73 ms for filler sentences as indicated in Table 3. In order to see whether there were any significant differences between the reaction times among these sentence types, a within subject paired *t*-test was conducted. Between the scrambled and canonical sentences, t (70)=3.68, $p <$ 0.001 suggesting that there was a significant difference in the RTs. Between the scrambled and filler sentences, t (70)=9.934, $p < 0.001$ also suggesting that there was a significant difference between the RTs among these two sentence types. Lastly, between the canonical and filler sentences, t (70)=7.90, $p < 0.001$, again suggesting that these two sentence types had a significant difference in their RTs. The results are indicated in Table 6.

Table 6. *Paired Sample t-Test Among Sentence Types on Reaction Time*

	Mean	*SD*	*T*	df	Sig. (2-tailed)
Scrambled – Canonical	276.93	634.23	3.679	70	.000*
Scrambled – Filler	830.81	704.70	9.934	70	.000*
Canonical – Filler	553.88	591.03	7.90	70	.000*

Note. * indicates $p < 0.001$.

From these results there was a significant slowdown in processing scrambled sentences compared with canonical and filler sentences. Also, canonical sentences were processed slower compared with filler sentences. The RT data showed that comprehending scrambled sentences incur processing cost.

Chapter 6: Discussion

6.1 Research Questions and Hypotheses

Research question 1 asked how L2 Japanese learners' accuracy rate compares when processing canonical versus scrambled sentences in a listening task. The results showed that the accuracy rate of scrambled sentences is significantly lower than that for canonical sentences. In addition, many participants had significantly slower reaction times when listening to scrambled sentences compared with canonical sentences. Therefore, combining these results, participants not only struggle to comprehend case marking particles accurately but also incur processing cost when listening to scrambled sentences.

These results contradict the findings from Mitsugi and MacWhinney (2010) and Smith (2016). Although their studies used self-paced reading, in their results, there was no slowdown in the reading time as well as no error rate difference among scrambled and canonical sentences. This may be because their experiments had fewer participants and items and thus may not have been sufficiently powered to see the full effect of scrambled sentences. Given that their studies focused on reading comprehension, the outcome may differ when tested in different modality such as in listening in the present study. However, the present study's data is consistent with that from Iwasaki (2003) and Shigenaga (2012) in which reading comprehension of scrambled and canonical sentences was the modality tested and thus, it cannot be conclude that the different outcomes are due to different tasks. The current study has a similar pattern as with the latter studies in which the evaluation of participants' listening comprehension accuracy rate dropped significantly for scrambled sentences.

Research question 2 asked how L2 Japanese learners' WM affects comprehension of scrambled sentences. The results showed that WM measured by WM Recall and WM Total from the Listening Span Task had little influence in the accuracy of scrambled sentences indicating that even if participants had higher WM scores, they had greater difficulty processing scrambled sentences accurately. Moreover, the interaction between WM and scrambled sentences showed that scrambled sentences had a negative impact and constrained participants' WM. This suggests that L2 learners cannot utilize case marking particles as effectively in a timed setting regardless of their WM capacity. This is also reflected in the Exit Survey which asked participants on what aspect they were focused when doing the Target Task. 45 participants (63%) indicated that they focused on particles suggesting that they knew that case marking particles are an important aspect in Japanese grammar, but this recognition was not sufficient to prevent errors with regard to comprehending scrambled sentences. A similar result was shown by Mitsugi and MacWhinney (2016) in which participants had knowledge about case marking particles but many could not utilize the case markers fully to predict upcoming linguistic items online. Moreover, when Japanese is spoken, sometimes there is no distinguishing between word boundaries and often forms one single prosodic unit (Warner & Arai, 2001) making listening to case marking particle more challenging. Because of this phonological camouflage, L2 Japanese learners may not be able to quickly differentiate which word is a noun nor be able to identify its corresponding case marking particle leading to a lower comprehension of scrambled sentences. Therefore, the processing of case marking particles is difficult even with higher WM capacity in a timed setting, such as in a listening task, and plays a smaller role in comprehending scrambled sentences.

Research question 3 asked how L2 Japanese learners' sound recognition ability influences comprehension of scrambled sentences for L2 Japanese. LLAMA-D was going to be used as a measurement of L2 learners' sound recognition ability. However, because LLAMA-D had a very low reliability, no analysis was conducted using it as the data would not have been sufficiently reliable. Therefore, this question could not be answered. One of the reasons why LLAMA-D may have had a very low reliability in this study is because this was conducted as the last task of the experiment. Since participants had to concentrate on listening in the previous tasks, they may have felt fatigue by the time they were doing the LLAMA-D task. Therefore, if the order of the task were different, the reliability of the test may have been higher. Moreover, this task did not have a practice phase and the software may have been more challenging to use without any practice. Although explained in detail both visually and orally, participants tended to struggle maneuvering the program. This difficulty may have added more psychological cost in which participants could not focus fully in the task. Although LLAMA-D had been shown high reliability in some studies (Saito, Suzukida, & Sun, 2018b), it also has been noted that because it has not been extensively standardized, the outcome may not be as accurate and reliable (Granena, 2013a); therefore, care is needed when using LLAMA-D task. Future research is needed to address a different task for evaluating how sound recognition ability influences comprehension in order to answer this question.

6.2 Discussion

Given the lack of any relationship found between WM and the processing of scrambled sentences, the role of participants' proficiency was analyzed. Participants' proficiency was measured through the SPOT Test Version B which showed that proficiency also had little effect in comprehending scrambled sentences—the probability of being accurate being 0.05. This suggests that even if participants had high proficiency in Japanese, it was still challenging to process case marking particles in scrambled sentences. Since both WM and proficiency did not play an important role in the comprehension of scrambled sentences, high accuracy participants' background information and their answers in the Exit Survey were further examined to evaluate why they were successful in processing scrambled sentences.

There were 7 participants who were able to comprehend over 14 scrambled sentences correctly out of 16 (87.5%). Details are described in Table 7 along with their WM Recall, WM Total, and SPOT scores along with their home language(s). Even among high accuracy participants, WM or proficiency was not exceptionally high. Although many had a perfect score in the SPOT test, not all have high proficiency. This is aligned with the findings that WM and proficiency do not play a significant role in the comprehension with scrambled sentences even among participants with high accuracy.

To examine whether home language(s) did not affect the comprehension of scrambled sentences, participants whose home language permitted flexible word order were examined as well. The 2 participants who spoke Korean at home (see Table 7) were the only ones who had a high accuracy on scrambled sentences. The 6

Table 7. *Participants with High Accuracy on the Scrambled Sentences and their WM and SPOT Scores*

Participants	Home Language(s)	Canonical (n = 16)	Scrambled (n = 16)	Filler (n = 32)	WM Recall (n = 48)	WM Total (n= 48)	SPOT (n = 60)
6	English & Korean	16	14	31	37	68.5	60
9	English & Korean	15	14	30	40	83	56
29	English	13	15	31	47	91	58
31	English & Chinese	15	14	31	38	81.5	59
36	English	16	14	31	47	86	60
54	Chinese	14	15	32	39	75	60
64	Chinese	15	14	30	32	65.5	60

others whose home languages permitted flexible word order spoke English and an additional language (2 Japanese, 2 Tagalog, 1 Slovak, and 1 Russian) and had an accuracy of between 0 and 11 sentences in comprehending scrambled sentences. Given the small number of participants, a statistical analysis was omitted; however, participants' home language—whether it has a flexible word order or not—does not descriptively appear to have affected participants' comprehension of scrambled sentences.

In addition, in examining the Exit Survey of participants with high accuracy in comprehending scrambled sentences, all indicated that they focused on the particles or the subject of the sentence while doing the task. Further, 3 participants explicitly stated that they knew about the Japanese flexible word order (Participant 9, 29, and 31). Interestingly, an additional 10 participants, who did not demonstrate high accuracy in comprehending scrambled sentences (accuracy ranged from 6 to 12 sentences) also responded that they were aware of Japanese flexible word order. Since there are differences in the accuracy among those who stated knowing about flexible word order but did not demonstrate high accuracy and of these participants, and 9 out of 10 participants have indicated that they were focusing on particles while doing the task, these participants may have realized that the word order was scrambled sometime during the task and changed their listening strategy. Looking at their responses throughout the task, there was a trend that these participants' accuracy for processing scrambled in the scrambled sentences increased during the task. Therefore, participants with high accuracy in comprehending scrambled sentences may have noticed the scrambled word order at a relatively early stage and concentrated on remembering the case marking particles. In order to notice that the case markings were switched, different types of individual differences, other than WM and proficiency, may have affected the processing of scrambled sentences. Further research is necessary to elucidate these individual differences.

From the quantitative and qualitative analyses, participants tended to have a difficult time comprehending scrambled sentence regardless of proficiency, WM, and language background. Even if they had the grammatical knowledge about word order flexibility, in a timed setting, they were not able to fully utilize it to comprehend the sentence accurately. Moreover, interestingly, 6 participants mentioned in their Exit Survey that Japanese is a SOV language and among them, few have indicated that they were puzzled when they heard the OSV scrambled sentence structure. This suggests that participants who are accustomed to the SOV word order transferred their knowledge in which they thought that the first noun was the subject. This is a similar result as previous studies done by Kilborn and Ito (1989) and Matessa and Anderson (2000) although they were focused on simple word-strings. This data may suggest that L2 Japanese learners transfer their L1 knowledge when comprehending sentences and pay less attention to case marking particles.

Moreover, there seemed to be a gap between participants' knowledge about case marking particles and word order flexibility and their processing of it in an online setting. The gap may have come from instructions that the learners are receiving—the majority of participants are students who are learning or learned Japanese though U.S. universities where grammar is taught explicitly. Looking at the materials they may have used, word order flexibility is introduced from relatively early on; however, learners may not be exposed to scrambled sentences enough to process them effectively. *Genki* (Banno, Ikeda, Ohno, Shinagawa, & Tokashiki, 2012), one of the most commonly used textbook in the U.S. schools (Rollins, 2015), mentions about word order flexibility in Chapter 3. Also, another textbook that many participants in this study used was *Japanese: The Spoken Language* (Jorden & Noda, 1987) has word order flexibility introduced in Lesson 6. Although participants may be taught about word order flexibility, they may have less opportunity to use and be exposed to it.

Given these findings and the fact that spoken Japanese utilizes a variety of sentence structures, there exists a need to modify the education of L2 learners appropriately. In order for learners to gain competency with flexible word orders, instruction that focuses on making learners practice form-meaning links—processing instruction (VanPatten, 2002)—may be necessary. In a meta-analysis that compared the effectiveness between processing instruction and production-based instruction, those who received processing instruction performed better than the production-based instruction group in receptive test, and both groups performed equally well in productive test (Shintani, 2015). For effectively implementing processing instruction, there are three steps in how to alter learners' processing mechanisms (DeKeyser & Prieto Botana, 2015). The first step is for learners to be explicitly be taught about the properties and rules of the target form. In the second step, learners practice the target form in order not to use the default processing and to avoid incorrect form-meaning mapping. For the final step, activities that makes learners actively process the target form and connect it to its function is used. These activities can be either referential or affective activities. For L2 Japanese learners to overcome natural processing pattern—SOV sentence structure—when comprehending scrambled sentences, implementing processing instruction in teaching may improve learners' comprehension and accuracy.

In addition to processing instruction, textbook dialogues and exercises can use a mixture of canonical and scrambled sentences. Moreover, in class, teachers could speak in a natural conversation and not restrict their sentences to be just canonical in structure but include scrambled sentences even from a beginner level. In order for learners to spontaneously and efficiently use the language, development of knowledge that can be retrieved and applied automatically in real-time setting is necessary (Jiang, 2007). With more exposure and practice to scrambled sentences in classroom and in teaching materials, L2 Japanese learners may develop efficient listening strategy to process case marking particles accurately. Since the goal of L2 learning and teaching is for the learners to be able to use the language—not just memorize grammatical knowledge—further research needs to take place in order to identify which specific teaching methods could be reformed to raise learners' competence in processing case marking particles.

6.3 Limitations

Some limitations exist in this study. First, only L2 Japanese learners were tested, and no NS Japanese was included as a control group. Therefore, it is not certain whether NS will have a similar or a different result in terms of accuracy and reaction time when comprehending scrambled sentences. Since there is not a mutual understanding of how NS process scrambled sentences (Kim et al., 2009; Mazuka et al., 2002; Mitsugi & MacWhinney, 2010, 2016; Miyamoto & Takahashi, 2002; Witzel & Witzel, 2016; Yamashita, 1997), testing the same materials with NS Japanese could have made this study more appealing.

In addition, the participants recruited were intermediate-level Japanese learners, who are enrolled in a 3rd year course or above. From the SPOT Test, many were clustered around the 100% accuracy rate. If there were higher proficiency level participants, these learners may have had higher accuracy rate in comprehending scrambled sentences. Because the sample size is relatively homogenous in terms of their proficiency, a wider range of participants should be recruited for further study.

Another limitation with this study was the use of LLAMA-D as a measurement of sound recognition ability. As mentioned earlier, the use of LLAMA-D has been identified as a measurement that may not be as accurate and reliable in a high-stakes environment (Granena, 2013a). Therefore, another measurement of sound recognition ability needs to be used in order to answer whether it plays a role in comprehending scrambled sentences. Moreover, there was a trend for those who explicitly stated that Japanese word order can be scrambled that their accuracy improved from midway in the listening task. This group of participants may have a higher pattern recognition or analytical ability. In the future, these types of individual differences can be used to examine the relationship between them and comprehending scrambled sentences.

Chapter 7: Conclusions

This is one of the first studies that looked at the relationship between individual differences and processing of Japanese case marking particles for L2 learners. Results showed that participants with both high WM and proficiency still had difficulty in comprehending scrambled sentences, and they could not process the case marking particles accurately in a timed setting. Although the results did not show any relations between WM capacity and the comprehension of scrambled sentences, different types of individual differences, such as pattern recognition or analytical ability, may have played a role.

Moreover, some participants knew about the flexibility of word order, but many failed to utilize this knowledge online. However, once they noticed that the word order was switched, there was a trend that they were able to change their listening strategy and subsequently improve their accuracy. In order for learners to process case marking particles efficiently, classroom instruction and teaching materials may need to be revised. Since case marking particles are an important aspect in Japanese grammar, future studies in the relationship between individual differences and the comprehension of scrambled sentences are necessary as is research on what teaching methods could be developed to effectively increase L2 learners' competence of case marking particles.

Appendices

Appendix A: Target Task Sentences

Canonical Sentences:

1. John-ga Takeshi-o suupaa-de mimashita.
   ‘John saw Takeshi at the supermarket.’
2. Mary-ga Yuko-o denwa-de damashimashita.
   ‘Mary fooled Yuko through the phone.’
3. Chris-ga Naomi-o kawa-de tasukemashita.
   ‘Chris saved Naomi at the river’
4. Sophie-ga Kenji-o genkan-de machimashita.
   ‘Sophie waited for Kenji at the entrance’
5. Takeshi-ga John-o gakkou-de ijimemashita.
   ‘Takeshi bullied John at school.’
6. Yuko-ga Mary-o depaato-de sagashimashita.
   ‘Yuko looked for Mary at the department store.’
7. Naomi-ga Chris-o ashi-de kerimashita.
   ‘Naomi kicked Chris with her feet.’
8. Kenji-ga Sophie-o sakkaageemu-de ouen shimashita.
   ‘Kenji cheered for Sophie at the soccer game.’
9. John-ga Yuko-o naifu-de sashimashita.
   ‘John stabbed Yuko with a knife.’
10. Mary-ga Naomi-o kyoshitsu-ni annai shimashita.
    ‘Mary guided Naomi to the classroom.’
11. Chris-ga Kenji-o shashin-ni torimashita.
    ‘Chris took a photo of Kenji.’
12. Sophie-ga Takeshi-o konbini-ni oikakemashita.
    ‘Sophie chased Takeshi to the convenience store.’
13. Takeshi-ga Mary-o ie-ni tomemashita.
    ‘Takeshi gave shelter to Mary at his home.’
14. Yuko-ga Chris-o konsaato-ni tsurete ikimashita.
    ‘Yuko took Chris to the concert.’
15. Naomi-ga Sophie-o riidaa-ni kimemashita.
    ‘Naomi chose Sophie as a leader.’
16. Kenji-ga Chris-o kazoku ni shoukai shimashita.
    ‘Kenji introduced Chris to his family.’
17. Kenji-ga John-o mono-de nagurimashita.
    ‘Kenji hit John with an object.’
18. John-ga Naomi-o makura-de tatakimashita.
    ‘John hit Naomi with a pillow.’
19. Mary-ga Kenji-o ginkoo-de mikakemashita.
    ‘Mary saw Kenji at the bank.’
20. Chris-ga Takeshi-o meeru-de mushi shimashita.
    ‘Chris ignored Takeshi through email.’

21. Sophie-ga Yuko-o paatii-de shirimashita.
    ‘Sophie became acquainted with Yuko at the party.’
22. Takeshi-ga Chris-o kaisha-de homemashita.
    ‘Takeshi praised Chris at the company.’
23. Yuko-ga Sophie-o kouen-de tetsudai mashita.
    ‘Yuko helped Sophie at the park.’
24. Naomi-ga John-o tegami-de bakani shimashita.
    ‘Naomi made fun of John through a letter’
25. Kenji-ga Mary-o shigoto-de hagemashimashita.
    ‘Kenji encouraged Mary at work’
26. John-ga Kenji-o byouin-ni yobimashita.
    ‘John called Kenji to the hospital.’
27. Mary-ga Takeshi-o deeto-ni sasoimashita.
    ‘Mary invited Takeshi for a date.’
28. Chris-ga Yuko-o eki-ni mukaeni ikimashita.
    ‘Chris picked Yuko up at the station.’
29. Sophie-ga Naomi-o hoteru-ni kaeshimashita.
    ‘Sophie sent Naomi back to the hotel.’
30. Takeshi-ga Sophie-o jyugyouchuu-ni omoidashimashita.
    ‘Takeshi remembered Sophie during class.’
31. Yuko-ga John-o basutei-ni okurimashita.
    ‘Yuko dropped John at the bus station.’
32. Naomi-ga Mary-o kuukou-ni unten shimashita.
    ‘Naomi drove Mary to the airport.’

Note. The scrambled sentences were made by switching the order of the nominative NP and the accusative NP.

Filler Sentences:

1. John-ga hon-o toshokan-de karimashita.
   ‘John borrowed a book at the library.’
2. Mary-ga ongaku-o densha-de kikimashita.
   ‘Mary listened to music on the train.’
3. Chris-ga tokei-o furiimaaketto-de urimashita.
   ‘Chris sold a watch at the flea market.’
4. Sophie-ga saifu-o umi-de nakushimashita.
   ‘Sophie lost her wallet at the beach.’
5. Takeshi-ga hana-o doubutuen-de agemashita.
   ‘Takeshi gave a flower at the zoo.’
6. Yuko-ga poppukoon-o eigakan-de kaimashita.
   ‘Yuko bought a popcorn at the movie theater.’
7. Naomi-ga eigo-o kyanpu-de oshiemashita.
   ‘Naomi taught English at camp.’
8. Kenji-ga shinbun-o kafe-de yomimashita.
   ‘Kenji read the newspaper at a café.’
9. John-wa Kenji-yori se-ga takai desu.
   ‘John is taller than Kenji.’

10. Mary-wa Naomi-yori me-ga ookii desu.
    'Mary's eyes are bigger than Naomi's.'
11. Chris-wa Yuko-yori ashi-ga osoi desu.
    'Chris is faster than Yuko.'
12. Sophie-wa Takeshi-yori atama-ga ii desu.
    'Sophie is smarter than Takeshi.'
13. Takeshi-wa Chris-yori ryouri-ga jyouzu desu.
    'Takeshi is good at cooking than Chris.'
14. Yuko-wa Mary-yori kami-ga nagai desu.
    'Yuko's hair is longer than Mary's.'
15. Naomi-wa John-yori supootsu-ga suki desu.
    'Naomi likes sports more than John.'
16. Kenji-wa Sophie-yori ninki-ga arimasu.
    'Kenji is more popular than Sophie.'
17. John-wa takushii-de raamenya-ni tukimashita.
    'John arrived at the ramen shop by taxi.'
18. Mary-wa basu-de bijyutsukan-ni ikimashita.
    'Mary went to the museum by bus.'
19. Chris-wa hikooki-de apaato-ni kaerimashita.
    'Chris returned to his apartment by plane.'
20. Sophie-wa kuruma-de jyugyou-ni kimashita.
    'Sophie came to class by car.'
21. Takeshi-wa gosai-kara amerika-ni sundeimashita.
    'Takeshi is living in the US from age 5.'
22. Yuko-wa daigaku-kara chikatetsu-ni norimashita.
    'Yuko rode the subway from university.'
23. Naomi-wa mado-kara heya-ni hairimashita.
    'Naomi entered the room from the window.'
24. Kenji-wa hachiji-kara ofisu-ni imashita.
    'Kenji was at the office from 8 o'clock.'
25. John-wa atarashii kutsu-o hakimashita.
    'John put on the new shoes.'
26. Mary-wa furui jitensha-o moraimashita.
    'Mary received an old bicycle.'
27. Chris-wa omoi kaban-o mochimashita.
    'Chris carried a heavy bag.'
28. Sophie-wa kireina kimono-o kimashita.
    'Sophie wore a beautiful kimono.'
29. Takeshi-wa nigiyakana machi-o arukimashita.
    'Takeshi walked through a lively town.'
30. Yuko-wa nagai repooto-o kakimashita.
    'Yuko wrote a long report.'
31. Naomi-wa omoshiroi bideo-o tsukurimashita.
    'Naomi made a funny video.'
32. Kenji-wa oishii koohii-o nomimashita.
    'Kenji drank a good coffee.'

Appendix B: Participant Background Questionnaire

Participant Background Information for Japanese Non-native Speakers

Age: __________ Male __ Female __ Other __

Japanese Course Level: _______________ JLPT Level: ____________

1. Do you have any problem with hearing or sight? If yes, please indicate:

   ____

2. What language(s) do you speak at home? __________________________________
3. At what age did you start learning Japanese? ________________
4. How long have you been studying/did you study Japanese in college? ____(years)
5. Have you studied Japanese before entering college? If yes, where and for how long? ____
6. Have you studied other languages beside Japanese? If yes, which and for how long?____
7. Have you ever been to Japan? If yes, when and for how long?

   ____

8. Rate your own Japanese proficiency on the following scale by checking the numbers:

I can…	With Difficulty				Easily
Maintain a conversation with my friends in Japanese.	1	2	3	4	5
Maintain a conversation with my professor/supervisor in Japanese.	1	2	3	4	5
Discuss about social issues in Japanese.	1	2	3	4	5
Watch Japanese TV/movies without sub-titles.	1	2	3	4	5
Read short blog articles in Japanese.	1	2	3	4	5
Read newspaper articles in Japanese.	1	2	3	4	5
Write emails to my professor/supervisor in Japanese.	1	2	3	4	5
Write a short essay in Japanese.	1	2	3	4	5

9. If your native language is other than English, please answer the following questions.

TOEFL score: ___________ TOEIC score: ___________

Other English Test: ____

At what age did you start learning English? _______________

Where did you learn English? ____

How long have you been living in an English-speaking country? __________ (years)

Please rate your own **English** proficiency on the following scale by checking the numbers:

I can…	With Difficulty				Easily
Maintain a conversation with my friends in English.	1	2	3	4	5
Maintain a conversation with my professor/supervisor in English.	1	2	3	4	5
Discuss about social issues in English.	1	2	3	4	5
Watch English TV/movies without sub-titles.	1	2	3	4	5
Read short blog articles in English.	1	2	3	4	5
Read newspaper articles in English.	1	2	3	4	5
Write emails to my professor/supervisor in English.	1	2	3	4	5
Write a short essay in English.	1	2	3	4	5

Please use the space provided below this sheet to ask any questions or express any concerns about your participation that you may have at this time.

Appendix C: Exit Survey

Exit Survey

Thank you so much for your participation! Before ending, I would like you to take about 5 minutes to respond to a survey. Your responses to the survey are very important to the research question in this study. Please answer to the questions with your complete honestly. Your responses will be kept completely anonymous and used only for the purpose of this study.

1. Did you focus on any aspect(s) when you were listening to Japanese sentences in the **Listening Comprehension Task** (the first listening test)?

____
____
____
____

2. Did you focus on any aspect(s) when you were listening to Japanese sentences in the **SPOT Test** (the second listening test)?

____
____
____
____

3. Do you have any tips/methods on how to improve your listening skill in Japanese in general?

____
____
____
____

4. Do you know anything about Japanese sentence structure? Where did you learn/hear about it?

____
____
____
____

Thank you!

Bibliography

Artieda, G., & Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences*, *50*, 42–48. https://doi.org/10.1016/j.lindif.2016.06.023

Baddeley, A. (1986). *Working Memory*. New York: Oxford University Press.

Baddeley, A. (2003a). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2003b). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Banno, E., Ikeda, Y., Ohno, Y., Shinagawa, C., & Tokashiki, K. (2012). *Genki: An Integrated Course in Elementary Japanese* (2nd ed.). Tokyo: The Japane Times, Ltd.

Carroll, J. B., & Sapon, S. M. (1959). Modern language aptitude test. San Antonio, TX, US: Psychological Corporation.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. https://doi.org/10.3758/BF03214546

DeKeyser, R., & Prieto Botana, G. (2015). The Effectiveness of Processing Instruction in L2 Grammar Acquisition: A Narrative Review. *Applied Linguistics*, *36*(3), 290–305. https://doi.org/10.1093/applin/amu071

Doi, T., & Yoshioka, K. (1988). Speech processing constraints on the acquisition of Japanese particles: Applying the Pienemann-Johnston Model to Japanese as a

Second Language. *Proceedings of the 1st Conference on Second Language Acquisition and Teaching*, 23–33.

Forsberg Lundell, F., & Sandgren, M. (2013). High-level proficiency in late L2 acquisition: Relationship between collocational production, language aptitude and personality. In G. Granena & M. H. Long (Eds.), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* (pp. 247–272). Philadelphia: John Benjamins Publishing Company.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers*, *35*(1), 116–124. https://doi.org/10.3758/BF03195503

Gathercole, S. E. (2006). Commentaries on Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 545–598.

Granena, G. (2013a). Cognitive aptitudes for L2 learning and the LLAMA aptitude test: What aptitude does the LLAMA test measure? In G. Granena & M. H. Long (Eds.), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* (pp. 105–129). Philadelphia: John Benjamins Publishing Company.

Granena, G. (2013b). Individual Differences in Sequence Learning Ability and Second Language Acquisition in Early Childhood and Adulthood. *Language Learning*, *63*(4), 665–703. https://doi.org/10.1111/lang.12018

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, *29*(3), 311–343. https://doi.org/10.1177/0267658312461497

Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*, *36*(4), 611–645. https://doi.org/10.1017/S0272263113000727

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences*, *38*, 44–53. https://doi.org/10.1016/j.lindif.2015.01.019

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and the L2 reading skill. *Studies in Second Language Acquisition*, *14*(1), 25–38.

Hayashibe, H. (1975). Word order and particles: A developmental study in Japanese. *Descriptive and Applied Linguistics: Bulletin of the ICU Summer Institute in Linguistics*, *8*, 1–18.

Hinds, J. (1982). Case marking in Japanese. *Linguistics*, *20*, 541–557.

Ito, T., Tahara, T., & Park, W. (1989). Japanese particles as case markers: Developmental psycholinguistic comparison between Japanese and Korean, and within Japanese [in Japanese]. In & T. W. F. C. Peng, V. M. Peng, H. Miyake, M. Sasaki (Ed.), *Language in the individual and society* (pp. 93–111). Hiroshima, Japan: Bunka Hyoro.

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 Publishing.

Jiang, N. (2007). Selective integration of Linguistic knowledge in adult second language learning. *Language Learning*, *57*(1), 1–33. https://doi.org/10.1111/j.1467-9922.2007.00397.x

Jorden, E. H., & Noda, M. (1987). *Japanese: The Spoken Language Part 1*. New Haven and London: Yale University Press.

Khan, M. S., & Bryce, M. (2005). Beyond right and wrong: An innovative approach to teaching Japanese particles using animation. In B. Bourke (Ed.), *Innovative practice in Japanese language education* (pp. 1–16). Brisbane: Queensland Univeristy of Technology. https://doi.org/10.1016/j.jep.2011.11.008.An

Kilborn, K., & Ito, T. (1989). Sentence Processing Strategies in Adult Bilinguals. In B. MacWhinney & E. Bates (Eds.), *The Crosslinguistic Study of Sentence Processing* (pp. 256–291). 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*(2), 151–166. https://doi.org/10.1016/j.jneuroling.2008.07.005

Kobayashi, N. (2016). Japanese language proficiency assessment with the Simple Performance-Oriented Test (SPOT) as a primary focus. In M. Minami (Ed.), *Handbook of Japanese Applied Linguistics* (pp. 175–198). De Gruyter, Inc. https://doi.org/10.1515/9781614511830-011

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals1*, *22*(6), 529–540.

Koda, K. (1993). Transferred L1 Strategies and L2 Syntactic Structure in L2 Sentence Comprehension. *The Modern Language Journal*, *77*(4), 490–500.

Komori, S., & Banno, E. (1988). Shuudan tesuto ni yoru shokyuu bunpoo no shuutoku ni tsuite [The acquisition of elementary-level grammar through group test]. *Japanese Education*, 126–128.

Kuno, S. (1973). *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

Kuno, S. (1978). *Danwa no bunpoo [Discourse Grammar]*. Tokyo: Taishukan Shoten.

Leeser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning*, *57*(2), 229–270.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review*, *21*(4), 861–883. https://doi.org/10.3758/s13423-013-0565-2

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the Relationship Between Modified Output and Working Memory Capacity. *Language Learning*, *60*(3), 501–533. https://doi.org/10.1111/j.1467-9922.2010.00565.x

Mackey, A., & Sachs, R. (2012). Older Learners in SLA Research: A First Look at Working Memory, Feedback, and L2 Development. *Language Learning*, *62*(3), 704–740. https://doi.org/10.1111/j.1467-9922.2011.00649.x

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*(3), 379–413. https://doi.org/10.1017/S0272263112000125

Matessa, M., & Anderson, J. R. (2000). Modelling focused learning in role assignment. *Language and Cognitive Processes*, *15*(3), 263–292. https://doi.org/10.1080/016909600386057

Mazuka, R., Itoh, K., & Kondo, T. (2002). Costs of Scrambling in Japanese Sentence Processing (pp. 131–166).

Meara, P. (2005). *LLAMA Language Aptitude Tests The Manual*.

Miki, S. (2012). Working memory as a factor affecting L2 listening comprehension sub-skills. *Kumamoto University Studies in Social and Cultural Science*, *10*, 119–128.

Minai, U., Isobe, M., & Okabe, R. (2015). Acquisition and Use of Linguistic Knowledge: Scrambling in Child Japanese as a Test Case. *Journal of Psycholinguistic Research*, *44*(3), 287–307. https://doi.org/10.1007/s10936-014-9347-x

Mitsugi, S., & MacWhinney, B. (2010). Second language processing in Japanese scrambled sentences. *Research in Second Language Processing and Parsing*, *53*, 159–175. https://doi.org/10.2174/15734005126661609271505

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, *19*(01), 19–35. https://doi.org/10.1017/S1366728914000881

Mitsugi, S., MacWhinney, B., & Shirai, Y. (2005). Cue-Based Processing of Relative Clauses in L2 Japanese. *Processing*, 123–138.

Miyake, A., & Friedman, N. (1998). Indivudual Differences in Second Language Proficiency: Working Memory as Language Aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, N.J.: L. Erlbaum Associates.

Miyamoto, E. T., & Takahashi, S. (2002). Sources of Difficulty in Processing Scrambling in Japanese. In M. Nakayama (Ed.), *Sentence processing in East Asian languages* (pp. 167–188). Standford, California: CSLI Publications.

Miyamoto, E. T., & Takahashi, S. (2004). Filler-Gap Dependencies in the Processing of Scrambling in Japanese. *Language and Linguistics*, *5*(1), 153–166.

Mori, S. (2008). Individual Differences in L2 Acquisition of Japanese Particles. *Online Submission*, 1–25.

Murasugi, K., & Kawamura, T. (2005). On the acquistion of scrambling in Japanese. In J. Sabel & M. Saito (Eds.), *The free word order phenomenon: Its syntactic sources and diversity* (pp. 221–242). New York: Mouton de Gruyter.

Orita, N. (2017). Predicting Japanese scrambling in the wild, (i), 41–45.

Ortega, L. (2009). *Understanding Second Language Acquisition*. London: Hodder Education.

Otsu, Y. (1994). Early acquisition of scrambling in Japanese. *Language Acquisition Studies in Generative Grammar*, 253–264.

Oyama, H. (2010). Automatic Error Detection Method for Japanese Particles. *Polyglossia*, *18*, 55–63. Retrieved from http://en.apu.ac.jp/rcaps/uploads/fckeditor/publications/polyglossia/Polyglossia_V18_Oyama.pdf

Rogers, V., Meara, P., Aspinall, R., Fallon, L., Goss, T., Keey, E., & Thomas, R. (2016). Testing aptitude: Investigating Meara's (2005) LLAMA tests. *EUROSLA Yearbook*, *16*. https://doi.org/io.i075/eurosla.i6.07rog

Rollins, A. (2015). *The Japan Foundation Survery on Japanese Langauge Education Institutions 2015: U.S. Data*. Los Angeles. Retrieved from https://www.jflalc.org/ckfinder/userfiles/files/jle/JF_Survey_Report_2015.pdf

Saito, K. (2017). Effects of Sound, Vocabulary, and Grammar Learning Aptitude on Adult Second Language Speech Attainment in Foreign Language Classrooms. *Language Learning*, *67*(3), 665–693. https://doi.org/10.1111/lang.12244

Saito, K. (2019). The role of aptitude in second language segmental learning: The case of Japanese learners' English /ɹ/ pronunciation attainment in classroom settings. *Applied Psycholinguistics*, *40*(1), 183–204. https://doi.org/10.1017/S0142716418000528

Saito, K., Suzukida, Y., & Sun, H. (2018a). Aptitude, Experience, and Second Language Pronunciation Proficiency Development in Classroom Settings: a Longitudinal Study. *Studies in Second Language Acquisition*, 1–25. https://doi.org/10.1017/S0272263117000432

Saito, K., Suzukida, Y., & Sun, H. (2018b). Aptitude, Experience, and Second Language Pronunciation Proficiency Development in Classroom Settings: a Longitudinal Study. *Studies in Second Language Acquisition*, 1–25. https://doi.org/10.1017/S0272263117000432

Sano, K. (1977). An experimental study on the acquisition of Japanese simple sentences and cleft sentences. *Descriptive and Applied Linguistics: Bulletin of the ICU Summer Institute in Linguistics*, *10*, 213–233.

Sano, T. (1994). Early Acquisition of Copy & Movement in a Japanese OSV Sentence. *BUCLD Proceedings Supplement*, *8*, 253–264.

Sasaki, Y. (1991). English and Japanese interlanguage comprehension strategies : An analysis based on the competition model. *Applied Psycholinguistics*, *12*, 47–73.

Sasaki, Y. (1994). Paths of processing strategy transfers in Japanese and English as foreign languages: a competition model approach. *Studies in Second Language Acquisition*, *16*(1), 43–72.

Shigenaga, Y. (2012). Processing of Scrambled Sentences by Learners of Japanese as a Second Language. *Arizona Working Papers in SLA & Teaching*, *19*, 79–103.

Shintani, N. (2015). The Effectiveness of Processing Instruction and Production-based Instruction on L2 Grammar Acquisition: A Meta-Analysis. *Applied Linguistics*, *36*(3), 306–325. https://doi.org/10.1093/applin/amu067

Smith, M. (2016). L2 Learners and the Apparent Problem of Morphology: Evidence from L2 Japanese. In A. G. Benati & S. Yamashita (Eds.), *Theory, Research and Pedagogy in Learning and Teaching Japanese Grammar* (pp. 99–125). London: This Palgrave Macmillan. https://doi.org/10.1207/s153279

Suzuki, H., & Toutanova, K. (2006). Learning to Predict Case Markers in Japanese Hisami Suzuki. *Proceedings Ofthe 21st International Conference on Computational Linguistics and 44th Annual Meeting Ofthe ACL*, (July), 1049–1056.

Suzuki, S. (1977). Nihon no yoji ni okeru gojun horyaku [Lexical-ordering strategy in early Japanese children]. *The Japanese Journal of Educational Psychology*, *25*(3), 200–205.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, *52*(4), 755–803. https://doi.org/10.1111/1467-9922.00203

Warner, N., & Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *The Journal of the Acoustical Society of America*, *109*(3), 1144–1156. https://doi.org/10.1121/1.1344156

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology*, *49A*(1), 51–79. https://doi.org/10.2307/1883501

Williams, J. N. (2010). Initial incidental acquisition of word order regularities: Is it just sequence learning? *Language Learning*, *60*(SUPPL. 2), 221–244. https://doi.org/10.1111/j.1467-9922.2010.00607.x

Williams, J. N., & Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. *Lingua*, *118*(4), 522–553. https://doi.org/10.1016/j.lingua.2007.03.003

Witzel, J., & Witzel, N. (2016). Incremental sentence processing in Japanese: A maze investigation into scrambled and control sentences. *Journal of Psycholinguistic Research*, *45*(3), 475–505. https://doi.org/10.1007/s10936-015-9356-4

Yagi, K. (1992). The Accuracy order of Japanese Particles. *Japanese Language Education Around the Globe*, *2*, 15–25.

Yagi, K. (1996). Shokyuu gakushuusha no sakubun ni mirareru nihongo no joshi no seiyoo junjo [Acquisition order of Japanese particles found in beginning-level Japanese learners' compositions]. *Japanese-Language Education around the Globe*, *6*, 65–81.

Yamashita, H. (1997). The Effects of Word-Order and Case Marking Information on the Processing of Japanese. *Journal of Psycholinguistic Research*, *26*(2), 163–188. https://doi.org/10.1023/A:1025009615473

Yamashita, H. (2002). Scrambled sentences in Japanese: Linguistic properties and motivations for production. *Text*, *22*(4), 597–633.