THE EFFECTS OF OUTPUT-BASED FOCUS ON FORM ON JAPANESE EFL LEARNERS' IMPLICIT KNOWLEDGE DEVELOPMENT THROUGH A TEXT-RECONSTRUCTION STORY-RETELLING TASK

By

Kiyotaka Suga

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Teaching English to Speakers of Other Languages – Master of Arts

2019

ABSTRACT

THE EFFECTS OF OUTPUT-BASED FOCUS ON FORM ON JAPANESE EFL LEARNERS' IMPLICIT KNOWLEDGE DEVELOPMENT THROUGH A TEXT-RECONSTRUCTION STORY-RETELLING TASK

By

Kiyotaka Suga

This quasi-experimental study investigated the effects of a text-reconstruction output task called Story-Retelling (SR) on Japanese English as a foreign language (EFL) learners' implicit knowledge development on three types of English relative clauses (RCs). Thirty-three Japanese English as a Foreign Language (EFL) university students were assigned to one of the three instructional groups: the SR group, the explicit grammar explanation (EGE) group, and the comparison group. They went through three 60-minute instructional treatments and three testing sessions: pretest, immediate, and delayed posttests. Responding to a measuring issue of previous effect-of-instruction studies, two different tests were administered: an Oral Elicited Imitation (OEI) test for measuring implicit knowledge (Ellis, 2005; Erlam, 2006) and an Untimed Fill-in-the-Blank (FITB) test for measuring explicit knowledge. Differential effects of the instructional treatments depending on the degree of typological markedness and processing difficulty of the three RC types were also explored. Contrary to the original expectations, results of the OEI test revealed that the SR group failed to show unique effects of output practice on the learners' implicit knowledge development, compared to the test-performances demonstrated by both EGE group and comparison group. A careful re-examination of the research designs and cognitive processes that involve in the OEI test pointed out some important methodological directions for future studies regarding the use of the OEI test. On the other hand, results of the FITB test showed a clear advantage for the EGE group at both immediate and delayed posttests.

ACKNOWLEDGEMENTS

This thesis would not have been completed without the help of a number of people. I would like to express my sincere gratitude to them here.

My deepest gratitude goes first and foremost to my mentor, Professor Shawn Loewen, for guiding me to accomplish this thesis project. Without his generous guidance and support, this thesis would not have been completed. Throughout this project, he has always responded to my questions quickly and helped me design this study, analyze the data, interpret the results of the study. Moreover, I would also like to thank Professor Loewen for leading me into the field of Instructed Second Language Acquisition (ISLA). Since I attended his two-day intensive ISLA seminar at Temple University in Tokyo three years ago, his passion and work kept encouraging me to further study and explore numerous issues in ISLA, especially the issues related to implicit and explicit knowledge development and form-focused instruction.

I would like to thank Professor Patti Spinner for giving me numerous pieces of advice on this study. I gained a number of insights for the design, analysis, and interpretation of the results of the thesis from her comments and suggestions. Without her support, I would not have been able to conduct this study.

I would also like to thank Ryo Maie for editing and proof-reading my theis and giving me a number of helpful pieces of advice out of his very busy schedule.

My special thank also goes to Dr. Keiko Imura at Takushoku University for supporting this research project. She helped me recruit participants and arrange the entire processes of data collection at her university. As my former MA thesis adviser, she also gave me a number of

pieces of advice from the designing stage of the research project. Without her cooperation, this study could not have been conducted.

Needless to say, this research would not have been possible without the cooperation of the participants. I would like to express my deep appreciation to all of them for eagerly participating in the instructional treatments from early morning and late evening before and after their course work at the university. They also gave me a number of insights to teach English and content in a university classroom. I would also like to thank Takumi Sato for helping me pilot the materials, recruit the participants, and conduct the instructional treatments.

Lastly, I would like to express my genuine gratitude to my mother for her wholehearted support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Background of the Study

One of the most engaging questions in second language (L2) instruction is how to introduce form-focused instruction (FFI) in the classroom context to effectively promote L2 development (Izumi, 2013b; Spada & Lightbown, 2008). Previous Instructed Second Language Acquisition (ISLA) studies have shown that not only receiving 'comprehensible input,' as proposed by Krashen (1982), but also output produced by L2 learners, especially in meaningful contexts, plays crucial roles to facilitate the processes of language learning and then develop learners' linguistic knowledge (e.g., de Bot, 1996; Izumi, 2002a, 2003a; Izumi & Bigelow, 2000; Izumi & Izumi, 2004; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Kowal & Swain, 1994; Morgan-Short & Bowden, 2006; Muranoi, 2007b; Swain, 1985; 1995, 1998, 2005; Toth, 2006; Uggen, 2012; among others). In spite of the accumulation of these studies on output, what has not been fully investigated is what kind of output practice contributes to the development of L2 learners' proficiency, which enables them to use the language in a spontaneous communication (Morgan-Short & Bowden, 2006; Muranoi, 2007a). Muranoi (2007a) argues that reconstructing a text that learners have comprehended can be predicted to effectively promote the four functions of output (i.e, the noticing, hypothesis testing, metalinguistic, and automaticity functions), and then eventually contribute to the development of L2 proficiency (also see Izumi, 2002, 2003a; Muranoi, 2007b; Swain, 2005).

Along with the importance of output in language acquisition, the type of linguistic knowledge developed through certain types of FFI is also a key issue in L2 teaching as well as L2 research. To help L2 learners use their target language in a real communicative situation, it is

crucial to develop their implicit knowledge of the language, which can be used automatically without conscious monitoring of their own language use (Ellis, 2005, 2015). Although the development of implicit knowledge is a key in L2 acquisition, Doughty (2003, 2004) pointed out that the outcome of instruction in previous ISLA studies was not precisely measured through the use of valid measurement which can assess L2 learners' implicit knowledge development.

Aims of the Study

Based on the results and the limitations of the previous ISLA studies, the primary purpose of the present study was to investigate whether a text-reconstruction output task termed Story-Retelling (SR) introduced in the framework of focus on form (FonF), which is a type of instructional approach that aims at drawing learners' attention to certain linguistic forms within a predominantly meaning-focused communication (Long, 1991; Doughty & Williams, 1998), contributed to the development of Japanese EFL learners' implicit knowledge on English relative clauses (RCs). Responding to a measuring issue of previous effects-of-instruction studies stated above (e.g., Doughty, 2003, 2004; Norris & Ortega, 2000), this study employed two different tests (i.e., an Oral Elicited Imitation [OEI] test and an Untimed Fill-in-the-Blank [FITB] test) to separately measure the impacts of the instruction on implicit and explicit knowledge, respectively. Furthermore, the effects of SR were also compared with the effects of explicit grammar explanation, which is one of the most common L2 teaching practices in many EFL contexts and can be categorized as focus on forms (FonFs) in order to examine how two different instructional approaches (i.e., FonF and FonFs) differed in terms of the impacts on the two different types of L2 knowledge: implicit and explicit knowledge. In addition to the issue of types of FFI, this study also explored how the complexity of the target grammatical forms

interacts with effects of instruction by targeting the three types of RCs: Subject-type (SU), Direct-object-type (DO), and Object-of-preposition-type (OPREP).

Organization of the Study

This thesis consists of six chapters. Following this introduction chapter, Chapter 2 reviews previous literatures on issues on FonF, implicit and explicit knowledge development in L2 learning, output-based FonF, and grammatical complexity and L2 instruction. All of these issues motivated the researcher to address the research questions in this study. In Chapter 3, the methodological procedures of the present study are presented including the whole research design, details of the participants, the target forms, procedures for the instructional treatments, testing instruments, scoring procedures, and analyses of the data. In Chapter 4, the results of the present study are given. Descriptive statistics, their visuals, and results of inferential statistics are presented. Discussions on the results presented in Chapter 4 are reported in Chapter 5. The last chapter summarizes the major findings and describes limitations and pedagogical implications of the present research.

CHAPTER 2

LITERATURE REVIEW

Focus on Form (FonF) in L2 Instruction

FonF is an instructional approach that aims at drawing learner's attention to certain linguistic forms within a meaningful communication (Long, 1991; Doughty & Williams, 1998). This instructional approach was introduced to strike a balance between two opposite instructional approaches: traditional grammar-based teaching and extreme meaning-oriented communicative language teaching (CLT). Traditional grammar-based language teaching, which is often conducted by the combination of detailed rule-explanations and audio-lingual pattern drills, tends to focus only on certain target linguistic forms in less-meaningful, decontextualized manner. On the other hand, the main focus of the strong version of CLT is placed exclusively on the meaning that is conveyed in communication; hence, it is likely that learners pay little attention to the target-linguistic forms. Previous ISLA studies have reported that both of these opposite instructional approaches failed to develop learners' well-balanced communicative competence (Brown, 2000; Izumi, 2002b; Loewen, 2013; Long, 1998; Richards & Rogers, 2014). Thus, dealing with the limitations of these two instructional approaches in the L2 classroom while capturing each strength has been a long-standing challenge in many L2 teaching contexts. To overcome this challenge, FonF, which briefly and appropriately directs learners' attention to certain target-linguistic forms within meaning-focused L2 tasks, has a great potential (Izumi, 2002b).

Measuring Effect of FonF on Learners' Implicit Knowledge Development

One of the important questions in FonF research is what kind of knowledge, either explicit

or implicit knowledge, is developed through a particular type of FonF instruction (Izumi, 2013b). While explicit knowledge is the type of knowledge that is consciously available to learners when they have enough time for controlled processing (Ellis, 2005; Ellis, Loewen, Elder, Erlam, Philp & Reiders, 2009), implicit knowledge is not consciously available but can be used intuitively and automatically in real communication without conscious awareness (Ellis, 2005; Loewen, 2015). Because of the obvious advantage of implicit knowledge in real communication, Ellis (2015) claimed that the prime goal of L2 instruction should be developing learners' implicit knowledge. However, it is still debated whether grammar should be taught explicitly or implicitly to achieve the ultimate goal. Norris and Ortega's (2000) meta-analysis showed that explicit type of instructions led to higher effect size than implicit type of instructions. However, Doughty (2003, 2004) pointed out that most of the studies which were treated in Norris and Ortega's (2000) meta-analysis were biased because the research period of these studies was very short and the measuring tasks were often decontextualized, which is advantageous for measuring learners' explicit knowledge. This methodological issue needs to be further addressed by empirical studies that employ a type of measurement that can accurately assess L2 learners' implicit knowledge development.

To address this methodological issue, Ellis (2005) investigated the reliability and validity of the following five measurements for explicit and implicit knowledge: an Oral Elicited Imitation (OEI) test, an Oral Narrative test, a Timed Grammaticality Judgement Test (TGJT), an Untimed Grammaticality Judgement Test (UGJT), and a Metalinguistic Knowledge test. The results of the study revealed that the first three tests were found to be reliable and valid measurements for implicit knowledge while the latter two tests were found to be better at measuring explicit knowledge. Particularly, he concluded that the OEI test can solidly measure

learners' implicit knowledge. Other similar studies also supported the conclusion of Ellis' (2005) study (e.g., Bowles, 2011; Erlam, 2006; Kim & Nam, 2017; Spada, Shiu, & Tomita, 2015, also see Rebuschat, 2013 for a methodological review on measuring implicit and explicit knowledge). Based on the findings from a range of psychometric validation studies, Ellis (2015) proposed that the following four criteria are the key-factors to create more valid measuring-tools for implicit and explicit knowledge: (1) degree of awareness (the use of intuition or rule knowledge), (2) time availability (timed or untimed), (3) focus of attention (meaning or form), and (4) degree of utility of knowledge of metalanguage. Although the OEI test has been supported as a valid measure of learners' implicit knowledge by these previous studies, a recent study challenged the findings of these studies, claiming that the OEI test does not actually measuring implicit knowledge but rather measuring another construct, which is automatized explicit knowledge (see Suzuki and Dekeyser, 2015; Suzuki 2017). However, as Dekeyser (2003, 2017) argues that these two types of knowledge are functionally equivalent, both of these two types of knowledge can be used in real-life, spontaneous communication, which is a primary goal for most L2 teaching (Ellis, 2015). Thus, even if the knowledge that the OEI test is measuring turns out to be not exactly implicit knowledge but rather highly proceduralized, explicit knowledge through the use of careful laboratory psycholinguistic experiment, the main focus of the current study is to explore whether L2 learners can still use the knowledge that is gained through form-focused instruction in their spontaneous communication. Moreover, the issue is still being investigated in several replication studies (Godfroid et al., 2018). Therefore, considering the number of replication studies that supported the findings of Ellis (2005) and the main goal of L2 teaching, the present study was conducted based on the four key criteria proposed by Ellis (2005, 2015).

Output-Based FonF and L2 Learning

In addition to the measurement issue of FonF research, effects of diverse types of FonF studies need to be further investigated (Izumi, 2013b), especially output-based FonF instructions that is conducted in meaningful L2 classroom contexts (Morgan-Short & Bowden, 2006; Erlam, et al., 2009). Muranoi (2007a) claims that "having learners reconstruct a text (story) that they have comprehended is one of the most effective instructional techniques that elicit learner output and eventually promote L2 learning" (p. 67). Based on the four primary functions of output: noticing, hypothesis-testing, metalinguistic and automaticity functions (see de Bot, 1996; Izumi, 2003a; Swain, 2005 for more comprehensive reviews of the four functions of output), Muranoi (2006) introduced a story-retelling (SR) task, which is a type of text-reconstruction task in which learners are asked to reconstruct a text that they have comprehended using a concept map. The concept map is a type of lexical representation which indirectly guides learners to use specific target-linguistic forms. Contrary to the dictogloss task (see Kowal & Swain, 1994; Swain, 1998; Swain & Lapkin, 1995), the first comprehension phase of the SR task does not involve listening and learners' self-notetaking. Since L2 learners tend to primary focus on lexical elements while engaging in production tasks (Swain & Lapkin, 1995; Hanaoka, 2007; Hanaoka & Izumi, 2012; Uggen, 2012), providing the lexical items using a concept map in the SR task reduces the cognitive load for retrieving these lexical items and enables learners to direct their attention to the target-grammatical forms more systematically, and guides them to use these forms during the retelling than does the dictogloss task, which often fails to direct learners' attention to specific target-grammatical forms.

Cognitive rationale of the SR task can also be explained by Levelt's (1989) Speech Production Model, which represents psycholinguistic mechanisms of speech production (see

Figure 1). This model can be divided into three major components: the Conceptualizer, the Formulator, and the Articulator. First, the starting point of speech production is a message created in the Conceptualizer. Then, this information is sent to the next component, the Formulator, in which the generated preverbal message is converted into a phonetic plan (internal speech) through two important processes: the grammatical encoding and phonological encoding. As presented in the figure, these two processes interact with lexical information stored in the mental lexicon. Finally, the actual speech is produced in the last component, the Articulator. In L1 production, the last two processes are subconscious and highly automatic processes (Levelt, 1989). However, this is not necessarily the case in L2 production, which can involve a great deal of controlled processing in the last two components (the Formulator and the Articulator) (Izumi, 2003a; Muranoi, 2007a). Therefore, it is crucial for L2 learners to systematically focus on practicing these processes. In the SR task, the lexical items and the messages to produce (or retell) have already been provided by the use of a concept map, thereby the learners can spend their limited cognitive resources for the processes of grammatical encoding, phonological encoding, and articulation while engaging in the SR task (Izumi, 2003a; Muranoi, 2007a).

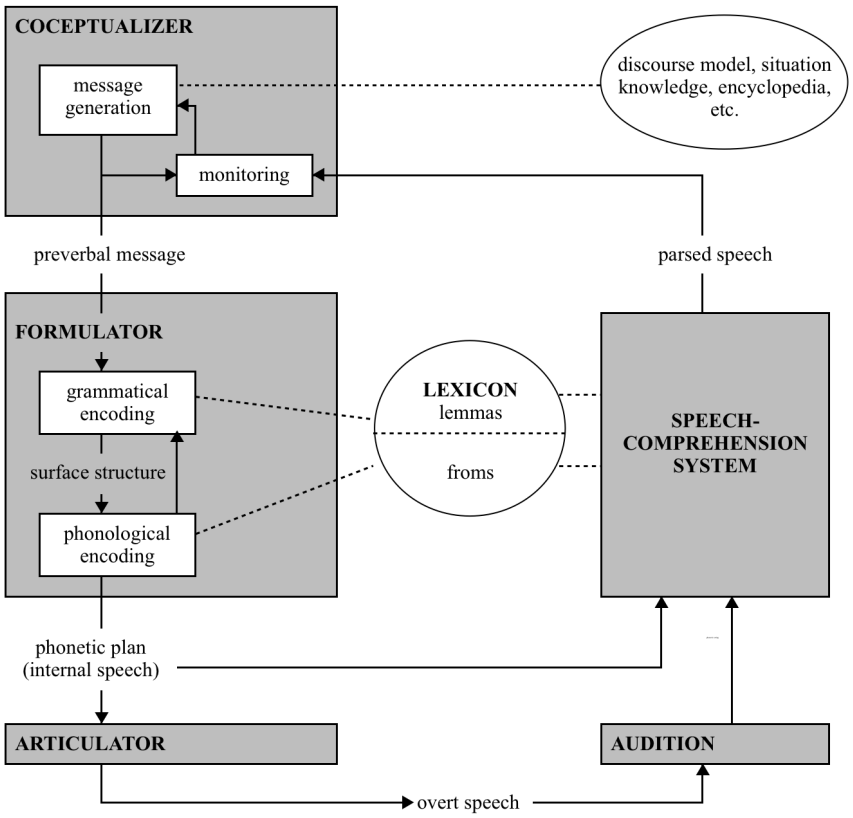


Figure 1. Levelt's Speech Production Model (from Levelt, 1989, p.9)

One study that investigated the effects of this type of text-reconstruction task is Muranoi (2007b), which examined the effects of text-summary-writing task termed *focus on form through guided summarizing* (FFGS). In FFGS, learners are asked to summarize a reading-text using a concept map. Contrary to the SR task, the FFGS task was conducted in a written mode, rather than an oral mode. During the task, the learners engage in written summary task twice. Between the first and second summarizing performances, an interval reflection-time was given to promote output-induced noticing by comparing their first summarizing-performance with the original text (i.e., cognitive comparison). The results of the study revealed that FFGS had a positive effect on learners' accuracy development in the use of English perfect passive and the effects of the task

held over two months. He attributed these positive effects of FFGS to the facilitative roles of output on learners' cognitive processes of L2 acquisition, such as output-induced-noticing, comparing, and hypothesis formulation and testing (see Gass, 1988, 1997 for detailed description on the cognitive processes of L2 acquisition form input to output). However, the study reported that the positive effects of FFGS were found only on the learners who were psycholinguistically ready to the target form. In other words, FFGS was not effective on the learners who had not had enough knowledge on the perfect active and the simple passive, both of which forms are considered as prerequisite to learn the perfect passive (Bardovi-Harling, 1995). Thus, it can be suggested that learners' acquisition stage or pre-instructional knowledge may be a crucial factor for an output-based FonF to be effective.

The impacts of output through a text-reconstruction task were investigated in a series of research by Izumi and his colleagues (Izumi, 2000, 2002a; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999).

Izumi (2002a) investigated whether a text-reconstruction task, an input enhancement technique or the combination of these two treatments promotes adult ESL learners' noticing and the acquisition of English relativization. In this study, the researcher compared the effects of instruction through four different measuring tasks: a sentence combination task, a picture-cued sentence completion task, an interpretation task, and a grammaticality judgment task. The study found that pushed output coupled with subsequent input-processing promoted learners' noticing and led to the acquisition of the target linguistic form (English relativization).

Based on the implications drawn from these previous text-reconstruction studies, the SR task can be predicted to promote noticing and contribute to the development of learners'L2 knowledge. However, paucity of empirical research actually investigated whether and how the

SR task, which is predicted to enable learners to systematically practice producing a target linguistic form in an oral mode, promote learners' L2 knowledge development. Moreover, most previous output studies that investigated the effects of text-reconstruction output were conducted in written mode (Izumi, 2002; Izumi & Bigelow, 2000; Izumi et al., 1999; Izumi & Izumi, 2004).

One study that tested the effects of the SR task is Suga (2016), which was a small-scale, exploring research that investigated the effectiveness of the SR task in an L2 classroom context. In this study, seven Japanese high school EFL learners went through four 90-minute instructional-treatment-sessions over one month. The effects were measured through two types of untimed oral- and written-fill-in-the-blank tests. The results of the study showed some positive effects on the learners' accuracy development on learners who received relatively high scores in the pretest. However, the instruction did not seem to have positive effects on the learners who received very low pretest-scores. One interesting finding from the follow-up interviews was that both in oral- and written-tests, the learners who improved their performances in the post-test phases worked on the tests relatively implicitly. One student commented that even though she could not verbalize the reason of her answer, she was able to work on the tests relatively smoothly, relying on her intuition. However, the biggest limitation of the study was the test-types used to measure the learners' performances. In the study, although both oral- and written-responses were elicited, the effects of the instructional treatment were measured only through an untimed fill-in-the-blank tests, both of which cannot precisely measure learners' implicit knowledge (Doughty, 2003, 2004). Thus, even though some learners' comments implied the development of their implicit knowledge, it was difficult to specify whether the SR task truly contributed to the development learners' implicit knowledge. Another methodological limitation of this study was that the factors that truly contributed to the learning gains were not specified

because of the small sample size, which made it difficult to employ a group comparison. Therefore, it was not be able to specify whether the gain was attributed to the first input exposure or the subsequent retelling of the story (output). These results and methodological limitations of the study generated a need for further research with a more rigorous research design.

The only output-study that employed precise measurements to assess explicit and implicit knowledge development is Erlam, Loewen, and Philp (2009), which compared the effects of meaning-oriented output instruction and processing instruction using an OEI test and an Untimed GJT. The results showed that both output instruction and processing instruction had positive effects on both implicit and explicit knowledge. This study revealed the facilitative effects of meaningful output on L2 learners' implicit knowledge development by the use of a more accurate measurement tools.

As reviewed in this section, producing output facilitates learners' noticing on the target-linguistic forms (e.g., Izumi, 2000, 2002a; Izumi & Bigelow, 2000; Izumi & Izumi, 2004; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Hanaoka, 2007; Russell, 2014; Song & Suh, 2008; Uggen, 2012) and may have some positive effects on the development of L2 knowledge. However, the effectiveness of output-based instruction is greatly influenced by various factors, such as learners' developmental stage on a targeted linguistic form (Muranoi, 2007b), task designs (Izumi & Izumi, 2004), and the degree of difficulty and complexity of the target forms (Izumi, 2002a, 2003b; Izumi & Bigelow, 2000).

Issues on Grammatical Complexity and L2 Instruction

Although it is quite challenging to determine the degree of difficulty of target linguistic forms (Dekeyser, 2005, 2016), there are several ways to make some predictions on it. One of the

most researched and supported predictions on the degree of linguistic-difficulty is the Noun Phrase Accessibility Hierarchy (NPAH), which hypothesized the universal order of linguistic difficulty of relativization from the most accessible type to the least accessible type based on the typological relationship of relative clauses (RCs) (Keenan & Comrie, 1977). The order predicted by the NPAH is: subject-type (SU), direct-object-type (DO), object-of-preposition-type (OPREP), genitive (GEN), and object-of-comparison (OCOMP). Another prediction that helps researchers to determine the degree of difficulties is Kuno's (1974) Perceptual Difficulty Hypothesis (PDH). Since effects of the matrix-position of the embedded clause was not considered in the NPAH, the PDH focused on learners' working memory capacity and claimed that any RC-types embedded in the subject position is more difficult to process than the RCs embedded in the object position, requiring more short-term memory capacity. One implicational study is Izumi (2003b), which investigated how the typological markedness and processing difficulty of English RCs interact with task-types (i.e., a sentence combination, interpretation, and grammaticality judgement task). Although this study revealed that L2 learners' performances on different English RC types were largely influenced by the combination of the place of embedded RC clause and task-demands (production or comprehension), the results of the study partially supported the prediction made based on the NPAH and suggested a complementary relationship of the NPAH and the PDH rather than a contradicting relationship. The findings of Izumi (2003b) allow detailed examination of learning effect in relation to typological complexity, processing difficulty of the target-linguistic form (i.e., English relative clauses), task types, and learners' developmental stages. However, few previous studies investigated effects of output on learners' implicit knowledge development in relation to the typological markedness and complexity of the target linguistic forms.

Motivated by the findings and limitations of the previous ISLA studies, the present study explored whether the SR task conducted in the framework of output-based FonF has positive effects on Japanese EFL learners' implicit knowledge development on English relative clauses using a type of test that can accurately measure learners' implicit knowledge development. To explore the possible advantages and limitations of both output-based FonF and explanation-based FonFs, the effects of the SR task were also compared with the provision of explicit grammar explanation. Finally, the effects of these two types of form-focused instruction (the SR task and the provision of explicit grammar explanation) were investigated in relation to the degree of the typological complexity of three types of relative clauses (i.e., RC[SU], RC[DO], and RC[OPREP]).

Research Questions and Hypotheses

The research questions addressed in the present study were as follows:

1. Does Story-Retelling (SR) make a unique contribution to developing Japanese EFL learners' implicit/explicit knowledge when it is compared with the provision of explicit grammar explanation?
2. Do the SR and explicit grammar explanation have differential effects on the development of Japanese EFL learners' implicit/explicit knowledge depending on the complexity of the target linguistic forms: RC (SU), RC (DO), and RC (OPREP)?

The hypotheses of the present studies are:

Hypothesis 1. As shown in the previous output studies, the SR task is hypothesized to have positive effects on Japanese EFL learners' performances on both the OEI test and the Untimed

Fill-in-the-Blank test. Also, the effects facilitated by the SR task hold over until the two-week delayed post-test (Muranoi, 2007b).

Hypothesis 2. Explicit grammar explanation contributes to Japanese EFL learners' performances only on the untimed fill-in-the-blank test in the immediate posttest, but the positive effects of the explicit grammar explanation do not hold over until the delayed post-test. Even though the provision of explicit knowledge through rule-explanations helps learners to understand specific grammatical features (Erlam, 2013), the effects of explicit instruction usually disappear very quickly (Doughty, 2004).

Hypothesis 3. The SR contributes to Japanese EFL learners' implicit knowledge development on all the three types of English RCs in both immediate and delayed posttests, while explicit grammar instruction has significant positive effects on their explicit-knowledge-development only on the simpler and less-marked RC types: RC (SU) and RC (DO), but not on RC (OPREP) in the immediate posttest. Dekeyser (2005) argues that implicit type of instruction is advantageous for complex grammatical rules while explicit instruction is good for simple rules. Although it is difficult to determine which grammatical forms are more complex and which forms are not (see Ellis, 2006; Dekeyser, 2005, 2016; Spada & Tomita, 2010), the present study decided the degree of difficulty of the three RC-types based both on the typological markedness predicted by Keenan & Cormrie's (1977) NPAH and processing difficulties of the RC-types reported by Izumi (2003b).

CHAPTER 3

METHOD

In this quasi-experimental study, two types of form-focused instructions (i.e., story-retelling [SR] and explicit grammar explanation [EGE]) were implemented to 33 Japanese EFL university students. The effects of these two types of form-focused instructions were compared with an exposure only control group (CG) through a pretest, immediate post-test, and delayed post-test (see Figure 2 for a summary of the overall experimental and instructional design). To examine whether the SR task and the explicit grammar explanation have differential effects on two types of linguistic knowledge (i.e., implicit knowledge and explicit knowledge), an oral elicited imitation (OEI) test and an untimed fill-in-the-blank test were used. The target linguistic forms were three types of English relative clauses (RCs): RC subject type (SU), RC direct-object type (DO), and RC object-of-preposition type (OPREP). The results of the study were analyzed quantitatively both with descriptive and inferential statistics. To explore the participants' background information and to supplement the results of these two tests, the participants completed an exit-questionnaire after the delayed posttest.

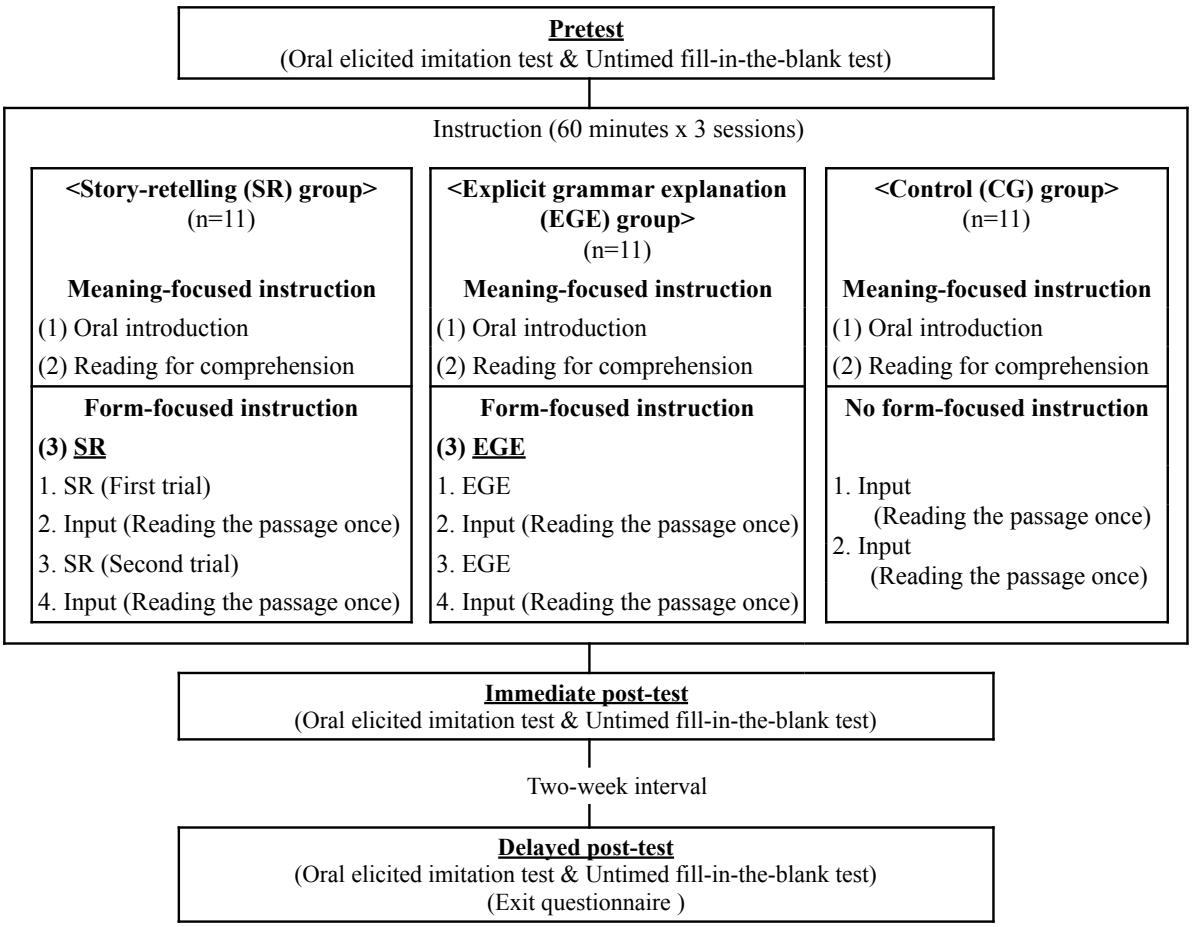


Figure 2. Experimental and instructional design

Participants

The participants of the study were Japanese EFL university students majoring in foreign language studies at a private university in Tokyo. Thirty-three students (eleven students for each instructional group) were recruited through the researcher's class visits and flyers at the university. The purpose of the research project was also briefly explained in the class visits.

The proficiency level of the participants was intermediate, ranging between 575 to 850 on the Test of English for International Communication (TOEIC), $M = 712.27$, $SD = 73.71$. In this university, all students were required to take the TOEIC test regularly (i.e., at least once a year)

for an assessment and course-placement purposes. The participants were equally distributed based on their TOEIC score (see Table 1 for the descriptive data of each group's TOEIC scores). After confirming all the assumptions were met, a one-way ANOVA was conducted to check if there were any statistically significant differences among the three groups (SR, EGE, and CG) in terms of their English proficiency (i.e., TOEIC scores). No significant differences were found between the results of each group's TOEIC scores, $F(2, 30) = 0.128$, $p = .88$, $\eta^2 = 0.008$. Due to scheduling constraints, participants in each group were not assigned based on their pretest performances (see the Results Section for the results of a one-way ANOVA on each group's pretest scores).

Table 1. *Participants' TOEIC Scores*

Group	*Mean*	*Min*	*Max*	*SD*	*Skew*	*Kurtosis*	*95% CIs*
SR	720.45	600	850	76.99	0.03	-1.30	[674.96, 765.94]
EGE	712.27	590	820	73.70	-0.14	-1.40	[668.72, 755.82]
CG	704.09	575	810	76.68	-0.22	-1.39	[658.77, 749.41]

Since the participants were recruited at a Japanese university, all of them had received formal instruction on the targeted linguistic forms (i.e., the three types of English relative clauses) in their previous formal education. Therefore, the target linguistic forms were not completely new to the participants. The proficiency factor and the amount of existing knowledge are very important in conducting an effect-of-instruction study. Previous ISLA studies have shown that partially acquired knowledge of the target linguistic form is a prerequisite condition to develop L2 learners' implicit knowledge through meaningful production practices (Doughty, 1991; Ellis, Loewen, & Erlam, 2006; Ellis & Shintani, 2014; Williams & Evans, 1998). Thus, the target population of the participants was a good candidate for measuring the impact of the SR task on learners' implicit knowledge development.

All the 33 participants did not miss any of the testing sessions and the instructional treatments. However, it was found that the recording of one participant's immediate post-test performance on the OEI test was not recorded successfully, and thus their entire test-performances on the OEI test from pretest to the delayed posttest were excluded from the analysis (see Table 8 in the Results Chapter). All the participants were asked to submit the consent form and their availability for the instructional treatments and testing sessions.

Target Forms

The impact of the instructional interventions was tested on the learners' L2 knowledge development for the three types of English restrictive relative clauses (RCs): RC subject-type (SU), RC direct-object-type (DO), and RC object of preposition-type (OPREP). The following are example sentences of these three types of English RCs embedded in both subject and object positions:

RCs embedded in the matrix subject position

RC (SU): *The woman who speaks Russian fluently* is my aunt.

RC (DO): *The car which the man drove* is very fast.

RC (OPREP): *The woman who(m) Bill is looking for* is beautiful.

RCs embedded in the matrix object position

RC (SU): The teacher liked *the girl who passed the exam easily*.

RC (DO): We like *the coat which Mary wears*.

RC (OPREP): She is *the woman who(m) I gave the book to.*

(Adopted from Izumi, 2003b, p. 288)

English RCs have been studied extensively in previous SLA studies (e.g., Ammar & Lightbown, 2005; Doughty, 1991; Hamilton, 1994; Izumi, 2002a, 2003b; Izumi & Izumi, 2004; Keenan & Cormrie, 1977; Kuno, 1974, Spada & Tomita, 2010). Since the current study attempted to explore the roles of output on L2 learners' implicit knowledge, these previous studies on English RCs would provide a rich source of information to compare with and analyze the impact of the instructional treatments on the participants' implicit- and explicit-knowledge development in relation to their psychological readiness, the formal complexity and processing difficulty of these three types of relative clauses. To carefully examine the results of the instructional treatments in relation to the degree of complexity and processing difficulty of the RC types, the three types of RCs (i.e., SU, DO, and OPREP), were chosen as the target linguistic forms.

Instructional Treatments

In the instructional treatments, each of the three treatment sessions was conducted in a different day within two weeks. All the instructional treatments were conducted by the researcher. As presented in Table 2, the pretest and the immediate posttest were conducted on the same day as the first and the third instructional treatments, respectively, due to the scheduling constraints. The data collection of this study was conducted during the summer break of the researcher, who is a graduate student of the US university. The research period was right before the end of the Spring semester of the Japanese university system. To complete all the instructional sessions and the testing sessions, the data collection had to be completed within a relatively short period following the time schedule presented in Table 2.

Table 2. *Experimental schedule*

First Session:	Pretest
	Instructional Session 1: Passage 1 (290 words)
Second Session:	Instructional Session 2: Passage 2 (319 words)
Third Session:	Instructional Session 3: Passage 3 (368 words)
	Immediate post-test
Fourth Session:	Delayed post-test (Two weeks after Day 3)
	Exit questionnaire

During the three 60-minute treatment sessions, three short passages consisting of about 300 words (M = 325.67, SD = 32.19) for each passage were introduced; each session used one of these passages. Each passage included two examples for each of the three types of RCs. For each of these two example sentences, one RC was embedded into the matrix of subject position and the other was embedded into the matrix of object position (see Figure 3 for the summary of the number of RC types included in the three reading passages).

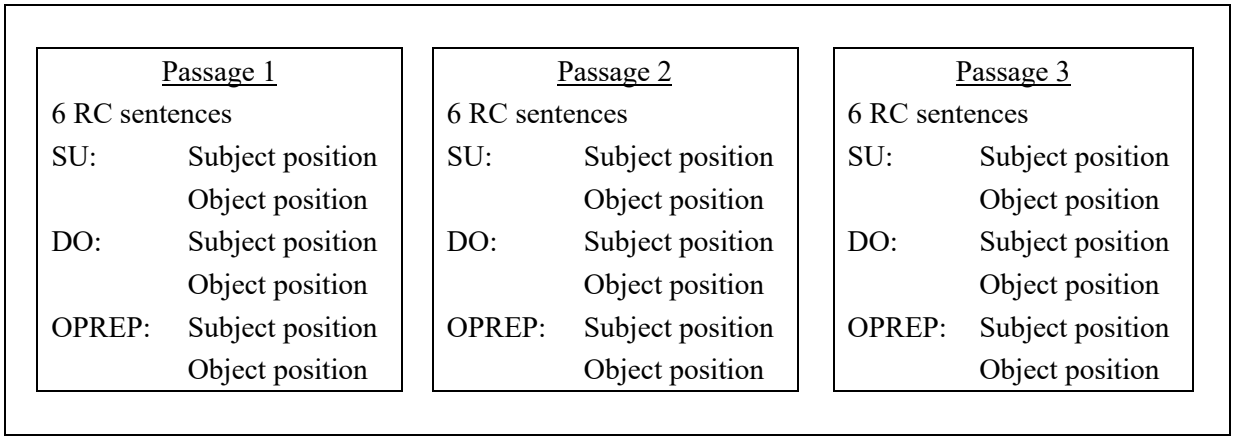


Figure 3. Summary of the RC types included in the three reading passages

All the three reading passages were written by the researcher considering the level of participants' English proficiency and the level of their vocabulary knowledge to maximize the participants' comprehension. Since previous vocabulary studies have shown that 95% of word coverage in reading text are necessary to attain a good comprehension of written texts (Laufer,

1989), 95% of all the reading passages consisted of words within the first three thousand frequent words (i.e., K1-K3 bands, Complete Lexical Tutor). Most of the words that were outside of the first three thousand frequent words were categorized as Japanese cognates (e.g., compass, baseball, scenario, passive, grammar, vocabulary, etc.), but vocabulary glosses of words outside the K1-K3 frequency bands were provided to each participant in the comprehension phase of each instructional treatment. In addition, the reading passage was read by two Japanese university students (one first-year and one second year university students), who were from the same program that the participants belonged to, in the pilot stage.

The content of the passages was decided based on the participants' interests and their major at the university so as to engage the participants in the meaning of the reading passages to benefit even the participants in the control group with some meaningful information just by participating in the study. Having talked with instructors at the university, it was found that most of the participants may have been interested in effective English learning and language learning in study abroad. For these reasons, the content of the passages was determined to be some topics related to SLA studies and English learning. Even though some basic concepts of SLA studies were introduced in the reading passages, any information that could potentially affect the participants' cognitive processes of L2 acquisition and the results of this research project was not provided (see Appendix A for all the reading passages used in the instructional treatments).

As shown in Figure 2, during the first phase of the instructional treatment, all three groups received the same meaning-focused instruction: (1) an oral introduction, in which the instructor provided background information of the reading passages (see Appendix B for detailed description and pre-reading questions asked in the first phase), and (2) a reading comprehension, in which the learners were asked to read a passage with vocabulary glosses (see Appendix C for

the vocabulary glosses used in the instructional treatments). After the first meaning-focused reading comprehension phase, each experimental group received a different form-focused instruction during the task phase (either the SR task, the explicit grammar explanation, or exposure only input).

Story-Retelling Task (Experimental Group 1, SR Group)

After the first oral introduction and reading comprehension phases, the participants of the SR group were instructed to engage in the SR task twice using a concept map (a schematic representation of keywords and phrases), which was aimed to indirectly guide learners to use the target-linguistic forms while reconstructing the text (see Figure 4; Appendix D). During the first and second trials of reconstructing the text, the learners were not allowed to access the original text. All the SR performances were audio-recorded to further examine whether each participant successfully used the target-linguistic forms during the SR performances. After performing the SR task each time, the participants were asked to read the original passage again. This input processing phase after the SR performance aimed to facilitate learners' subsequent output-induced noticing and cognitive comparisons, both of which were predicted to be promote learning of the target linguistic form.

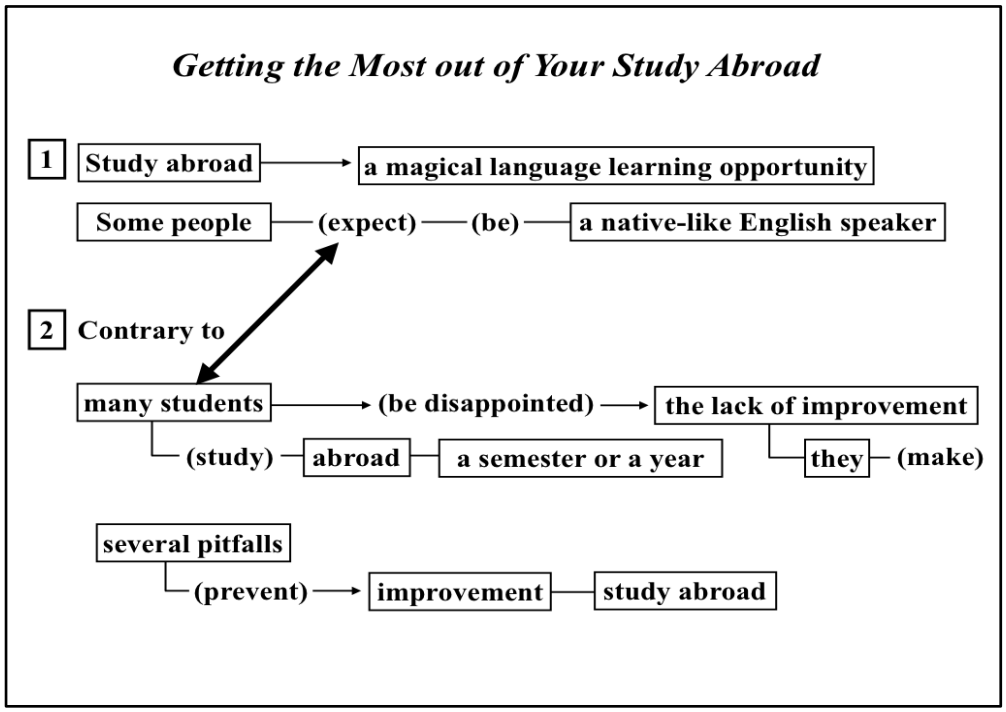


Figure 4. An example of a concept map used in the SR task

Explicit Grammar Explanation (Experimental Group 2, EGE Group)

The EGE group also went through the same instructional procedures until the task phase. After the comprehension phase, the EGE group received explicit grammar explanations on the target-linguistics forms in their first language (Japanese) using example sentences of RCs used in the reading passage. Since the SR group engaged in SR twice in the task phase, the explicit instruction group was also provided the explanation twice: (1) They received an oral-explanation by the researcher; and then (2) They were asked to read the summary of the explicit explanation given in Japanese on a handout (see Appendix E). The reason why the participants received the second explanation was to provide the equal amount of input-exposure as the SR group was exposed to. After receiving each grammar explanation, the participants were asked to read the passage once each time (twice in total during the task phase).

Exposure Only Comparison Group (Experimental Group 3, Comparison Group)

The comparison group only read the passage twice in the task phase without engaging in any additional form-focused task as did the other two experimental groups. In the task phase, both of the experimental groups were exposed to input after performing the SR task or receiving the explicit grammar explanation. Thus, the comparison group was also provided with the same amount of input: reading the passage twice in the task phase. In this study, the comparison group was not a test-only control group. Similar to the control group employed by Doughty (1991), the group was exposed to the target linguistic forms without receiving any form-focused instruction across all the instructional sessions. Thus, the comparison group of this study was an exposure only control group, which can be categorized as a focus on meaning (FonM) instructional condition.

Testing Instruments

The effects of the treatment were measured in three testing phases (pretest, immediate posttest, and delayed posttest) using two different types of measuring instrument: an OEI test, whose requirement of online production and the time pressure encourages the participants to use their implicit knowledge (Ellis, 2005; Erlam, 2006, 2009) and an untimed fill-in-the-blank (FITB) test, which allows learners to access their explicit knowledge with their conscious monitoring. One version of each test was used throughout the three testing phases, but the order of the statements was changed each time. The time-gap between the immediate posttest and delayed posttest was two weeks.

Pilot-testing for creating the testing instruments

To create these two testing instruments, two provisional tests were administered to a

diverse population including two native speakers of English, four Japanese EFL learners, three Japanese ESL learners attending a US university, and one highly-proficient international graduate student. The demographic information of the participants is summarized in Table 3.

Table 3. *The demographic information of the participants*

Students	Gender	Context	L1	English proficiency
S1	Female	Japanese adult EFL learner (EFL)	Japanese	TOEIC 650
S2	Male	Japanese university student (EFL)	Japanese	TOEIC 635
S3	Male	Japanese university student (EFL)	Japanese	TOEIC 465
S4	Female	Japanese high school student (EFL)	Japanese	N/A
S5	Female	Japanese high school student (EFL)	Japanese	Eiken 2nd
S6	Male	US university student (ESL)	Japanese	TOEIC 750
S7	Female	US university student (ESL)	Japanese	TOEFL iBT 52
S8	Female	US university student (ESL)	Japanese	TOEFL ITP 560
S9	Female	US graduate student (Linguistics)	Chinese	TOEFL iBT 109
S10	Male	US graduate student (Linguistics)	English	N/A
S11	Male	US graduate student (Non-Linguistics)	English	N/A

To examine the quality of the two provisional tests (i.e., the OEI test and the FITB test), item analysis was carried out by calculating item facility (IF) value, item discrimination (ID) value, and internal consistency. The results of the pilot tests are presented in Table 4. Based on the results of these calculations and comments from the follow-up interviews, which asked about (1) the pilot-tests in general, (2) unknown lexical items for the participants, (3) unclear instructions on both of the tests, and (4) too easy or too difficult items, all the provisional-test-items that were found to be problematic were revised. Furthermore, the overall research design was also revised based on the results of the pilot-tests.

Originally, the target population was Japanese EFL university students whose TOEIC scores were less than 500. However, the results of the pilot-test showed floor effects on the oral elicited imitation test, hence the target population was changed to the level of students whose

English proficiency (between 575 to 850 on TOEIC test) was higher than the original targeted-participants. Another revision was the format of the FITB test. To prevent ceiling effects observed in the pilot-test, the space of each blank was extended. While learners only needed to answer a relative pronoun and the modifying part, leaners were asked to answer all the elements in a relative clause, including an antecedent, which is a noun modified in a relative clause. The final version of both OEI and FITB tests are described in the following sections.

Table 4. *Descriptive statistics of the total scores on the provisional OEI and FITB tests*

Students	OEI Test	FITB Test
S1	3	24
S2	3	27
S3	0	16
S4	0	26
S5	2	28
S6	9	25
S7	3	27
S8	13	28
S9	26	N/A
S10	30	N/A
S11	29	N/A
N	11.0	8.0
Mean	10.7	25.1
Max	30.0	28.0
Min	0.0	16.0
SD	11.4	3.7

Highest possible score = 30

Oral Elicited Imitation (OEI) Test

The OEI test consisted of 24 statements, in which eight statements were created for each of the three types of RCs: SU, DO, and OPREP. In this study, no distractor items were used because the total duration of each OEI test had already been 12 minutes and it was not possible to spend more time for this test within the limited time for each testing session. Table 5

summarizes the number and the types of test-items tested in each testing phase.

Table 5. *The number and types of test-items in each testing phase*

	Pretest	Immediate posttest	Delayed posttest
Novel sentences	24	24	24
RC (SU)	8	8	8
RC (DO)	8	8	8
RC (OPREP)	8	8	8
Total	24	36	36

Informed by Erlam (2006) and Erlam et al. (2009), the OEI test of this study used a belief questionnaire, which directed learners' focus on the meaning of each statement. For this reason, the stimulus sentences were written in a way that the learners could judge whether the meaning of the sentences are true, false, or not sure. It has to be admitted that several sentences included wrong information or controversial information (see Appendix F for the stimulus sentences for the OEI test). The test was administered in the following procedures: (1) the learners listened to each statement; (2) the learners were asked to decide whether the statement was true, false, or whether they were not sure in four seconds; (3) after judging the content of each statement according to their beliefs, the learners were asked to quickly repeat the statement as accurately as possible in ten seconds (see Erlam et al., 2016 for detailed explanation on creating and administering an OEI test). Figure 5 presents sample slides used in the OEI test to direct the participants to take the test.

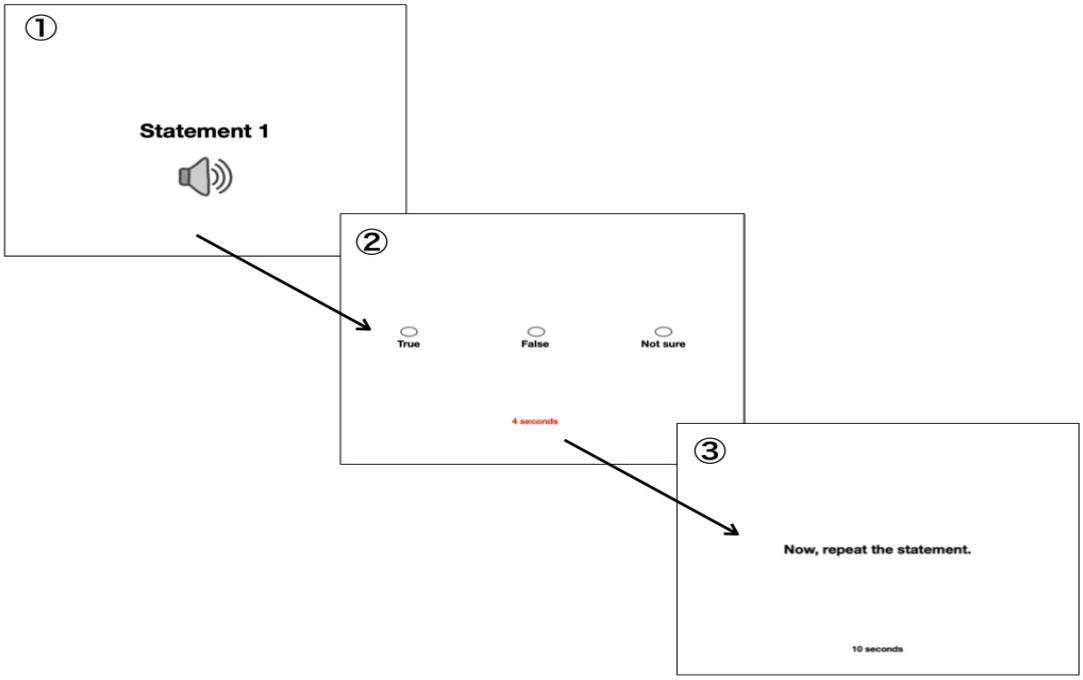


Figure 5. Sample slides used during the OEI test

Previous studies (e.g., Ellis, 2005; Erlam, 2006) included ungrammatical statements and encouraged learners to repeat the statement by correcting the ungrammatical parts. However, in this study, all the stimulus statements were grammatical because the target linguistic forms of the study were three types of RCs, which require syntactic processing. If the stimulus sentence were ungrammatical, it is very difficult to comprehend the meaning and, moreover, the learners may not be able to figure out which type of RC types should be used to repeat the statement.

Test reliability of the OEI test was measured using the Cronbach's alpha. A high coefficient of reliability was obtained in Pretest, Immediate posttest, and Delayed posttest, $\alpha = 0.90$, $\alpha = 0.89$, $\alpha = 0.88$, respectively.

Untimed Fill-in-the-Blank Test

The untimed fill-in-the-blank test was also created in the same structure as the OEI test (see Appendix F for examples of the test). All the test-items consisted of 24 items and eight

items were created for each RC type (see Table 5). To fill the blanks in this test, learners were provided with Japanese translation sentences and some vocabulary glosses for difficult words. In each test-sentence, an antecedent and the whole modifying clause were blanked. Example items of the FITB test are the followings:

1) **あなたの車を目的地まで案内してくれる最も良いアプリは、Google マップです。**(application ＝アプリ, navigate＝～を案内する, to your destination ＝目的地まで)
   - **(** **)** is *Google Map*.

   Correct Encoding:
   (The best application that navigates your car to your destination) is Google Map.

2) **たくさんのビーチがある 1 番の観光地は沖縄です。**
   (sightseeing place＝観光地、has a lot of beaches＝たくさんのビーチがある)
   - **(** **)** is Okinawa.

   Correct Encoding:
   (The best sightseeing place that has a lot of beautiful beaches) is Okinawa.

Again, using the Cronbach's alpha, test-reliability was examined on each testing session (Pretest, Immediate, and Delayed posttests), yielding a high reliability coefficient, $\alpha = .87$, $\alpha = .87$, $\alpha = .82$, respectively.

Scoring Procedures

Data gathered through the OEI test and the FITB test was calculated to quantify the participants' learning of the target linguistic forms (i.e., the three types of English relative clauses). In both tests, one point was given to a correct response and zero point was given to an

incorrect response.

First, to eliminate the participants who had already acquired the target linguistic forms prior to the instructional treatments, 80 percent accuracy or more in the pretest of the OEI test was arbitrarily set as the cut-off point. In addition, 20 percent or less in the pretest of the OEI test was also set as a cut-off point, which indicates almost no knowledge. However, no participants received higher or lower scores than the cut-off scores at the pretest stage.

Since the scoring greatly impacts the interpretation of the results, detailed scoring criteria were created to consistently score the participants' test-performances on the OEI test and the FITB test. The specific scoring criteria for each test are presented separately in the following sections.

Oral Elicited Imitation Test

To calculate the participants' test-performances, the correct use of the form and meaning were examined. Following the scoring criteria used in the previous studies (e.g., Erlam, 2006), if the participant responded exactly with the correct use of the target-linguistic form, the response was scored as a correct response. Even though the response was not exactly the same as the stimulus sentence, credit was also given to responses depending on the type of responses and the use of the target forms. The following types of sentences were considered as correct:

(1) The response included some different lexical items, but the meaning of the sentence was the same or almost the same as the stimulus sentence (e.g., using a different word but in the same category);

*Stimulus sentence: In Japan, there are many convenience stores that sell **vegetables**.*

*Response: In Japan, there are many convenience stores that sell **salad**.*

(2) The response that included some ungrammatical uses on linguistic forms other than the target-linguistic forms (e.g., incorrect uses on tense and aspect, the marker for third person singular –s, incorrect use of preposition, etc.);

Example 1

*Stimulus sentence: Ueno zoo is the first zoo that a panda **arrived** at.*

*Response: Ueno zoo is the first zoo that a panda **arrive** at.*

Example 2

*Stimulus sentence: Ichiro is a baseball player who **belongs** to an American team.*

*Response: Ichiro is a baseball player who **belong** to an American team.*

(3) The response that used OPREP for the elicitation of DO (but not for the other way around: DO for OPREP);

Correct (DO → OPREP)

*Stimulus sentence: Donald Trump is a person that most American people **respect**. (DO)*

*Response: Donald Trump is a person that most American people **respect for.** (OPREP)*

Incorrect (OPREP → DO)

*Stimulus sentence: Tokyo University is **the school that the Prime Minister Abe studied at**. (OPREP)*

*Response: Tokyo University is **the school that the Prime Minister Abe studied**. (DO)*

The reason why an elicitation of OPREP instead of DO was scored as correct but not the other way around was that RC (OPREP) is more marked-form than RC (DO) according to the NPAH

(Keenan & Cormrie, 1977), but both types of the RCs can be classified as the object-type RCs; thus, it may be predicted that the use of more marked form of the same type of RC may reflect the learners' correct use of the less-marked object-type RC. However, it is difficult to predict whether the learner can use a more marked form correctly based on an elicitation of a less-marked type of RC.

Next, the following elicitations were not credited and scored as an incorrect response;

(4) The response that was produced with no use of the target-linguistic form;

Stimulus sentence: President Obama is ***the first US president who visited Hiroshima****.*

Response: President Obama ***visited Hiroshima.***

(5) The response that was produced with incorrect use of the target-linguistic form;

Example 1

Stimulus sentence: In Japan, there are many convenience stores ***that sell vegetables****.*

Response: In Japan, there are many convenience stores ***sell vegetables****.*

Example 2

Stimulus sentence: Global warming is an issue that all Japanese people must think ***about****.*

Response: Global warming is an issue that all Japanese people must think.

(6) The meaning of the response was completely different from the stimulus sentence even though the form was correctly used;

Stimulus sentence: ***Ichiro*** *is a baseball* ***player who belongs to an American team****.*

Response: ***Major league*** *is a baseball* ***league that has many strong players****.*

(7) Even though the participants used the target-linguistic form correctly, if the sentence was incomplete (e.g., just providing the part of the relative clause), the response was coded as incorrect;

Stimulus sentence: A school subject that most elementary school students like ***is math****.*

Response: A school subject that most elementary school students like…

Untimed FITB Test

To calculate the score of the FITB test, the same scoring criteria used for the OEI test (see the scoring criteria from 1 to 7 for the OEI test) were employed. In addition to these scoring criteria, an additional criterion was used because, in contrast to the OEI test, there was no stimulus sentences in the FITB test. In this test, learners were asked to fill the blanks based on Japanese-translation sentences, more various types of elicitations were possible. The additional scoring criterion was:

(8) The elicitation that was written in the passive voice for an item of DO or OPREP was coded as incorrect;

Stimulus sentence: Oda Nobunaga is (a person ***that most Japanese people respect****).*

Response: Oda Nobunaga is (a person ***who is respected*** *by most Japanese people).*

If the passive voice was used for an item of RC(DO) or RC(OPREP), the answer was coded as incorrect, because object type RCs require learners to keep the information of the modified noun (the antecedent) in their working memory until they process the whole sentence. This is one of

the great difficulties related to the use of the DO type of RCs (Izumi, 2003b). If a learner substitutes a sentence of object type RC into passive voice, the quality of the sentence differs because the modifying clause of a passive RC clause starts with a be-verb, thus the learners do not need to keep the information of the modified noun until the position that noun was originally placed. In this study, all the test performances on both OEI and FITB tests were scored based on these eight scoring criteria described above.

Analyses

Analyses of the Effects of Instruction

The results from the two types of tests (the OEI and Untimed FITB tests) were analyzed quantitatively using both descriptive and inferential statistics. In order to analyze the participants' general English proficiency, one-way analyses of variance (ANOVA) was first conducted based on the participants' TOEIC scores for the group-assignment purpose. To examine the participants' pre-instructional knowledge level, the results of pretest scores on the OEI and FITB tests were also compared by performing a series of one-way ANOVAs.

Descriptive statistics of all the data from the pretest and the two posttests were first examined. Then, after checking required assumptions, the data were submitted to mixed-design analyses of variance (ANOVAs) using one between-subjects (Group) and one within-subjects (Time) factors. The between-subject factor (Group) had three levels: the SR group, the EGE group and the exposure only control group, and the within-subject factor (Time) also had three levels: Pretest, Immediate Posttest, and Delayed Posttest. The following sections, the results of the required assumption checking for the mixed design ANOVA on the results of the OEI and FITB tests are presented, respectively.

Assumption Checking for the Mixed-Design ANOVA on the Results of the OEI Test

First, the assumption of the normality was examined by visually inspecting the distribution of each histogram (see Figure 6). Eye-balling the shape of each histogram, none of them looked normally distributed, but most Shapriro-Wilk test-results did not show a significant result, except the result of the immediate posttest by the EGE group (see Table 6). The contradiction between the visual observation and the results of the Shapiro-Wilk test may have been due to the sensitivity of the Shapiro-Wilk test to sample sizes. Since the sample size of this study was small (10 or 11 learners for each instructional group), it may be better to conclude that the assumption of normal distribution was violated, prioritizing the descriptive information and the results of the visual observation (see Descriptive Statistics and its visuals presented in Table 8 and Figure 14, respectively in the next Result Chapter).

As for the variance of each dataset, the standard deviation values indicate relatively similar variance across the groups (see Table 8). Furthermore, the size of the boxes in the boxplots (see Figure 14) looked relatively similar among the three groups across the three testing sessions. These observations were confirmed by the results of Levene's test, which showed non-significant result, $F(8, 87) = 0.09$, $p = .10$. Thus, the equality of the variance of the datasets was satisfied.

Figure 7 and 8 show the distribution of the residuals extracted from the ANOVA model. As shown in Figure 7, the data points were not aligned along the straight line, indicating the distribution of the model residuals was not normal. The shape of the histogram of the residuals presented a bimodal distribution (see Figure 8). These observations were confirmed by the results of the Shapiro-Wilk test, $W = 0.96$, $p < .001$. The assumption of homogeneity of residuals seemed to be randomly scattered based on the visual inspection of Figure 9.

Finally, the assumption of sphericity was examined through the results of the Mauchly's sphericity test, which indicated a non-significant value, $W = 0.94$, $p = .44$, showing that the assumption of sphericity was satisfied.

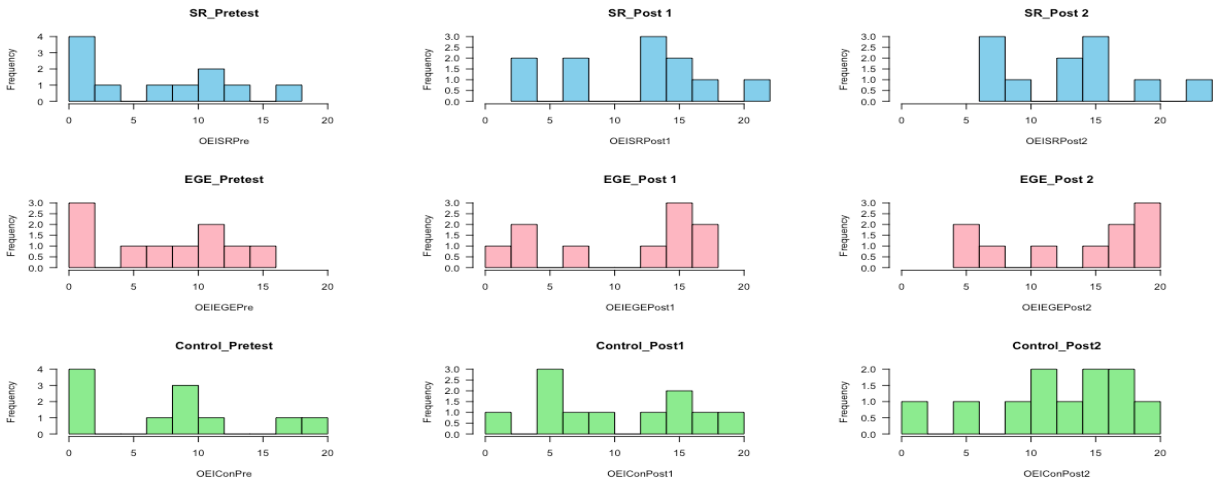


Figure 6. Histograms for the distribution of the OEI test scores

Table 6. *Results of Shapiro-Wilk test on the OEI test*

Group	Test	*W*	*n*	*p*-value
SR	Pretest	0.88	11	.11
	Posttest 1	0.94	11	.56
	Posttest 2	0.93	11	.42
EGE	Pretest	0.90	10	.23
	Posttest 1	0.82	10	.03
	Posttest 2	0.87	10	.11
Control	Pretest	0.92	11	.29
	Posttest 1	0.94	11	.51
	Posttest 2	0.94	11	.52

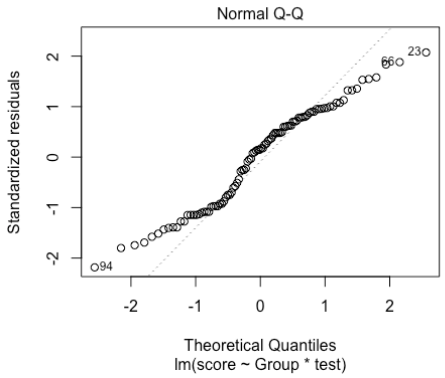


Figure 7. A normal qqplot for the residuals (OEI)

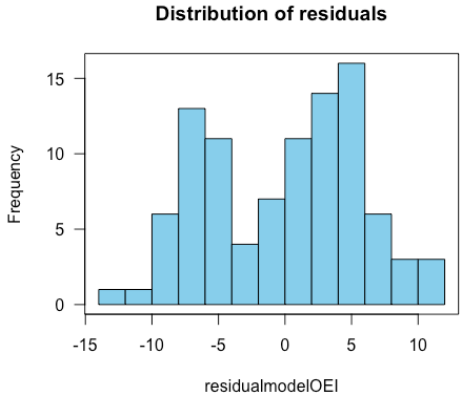


Figure 8. Histogram for the residuals (OEI)

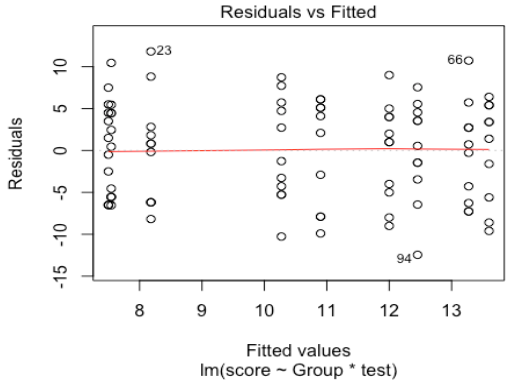


Figure 9. Residuals vs Fitted plots for the residuals (OEI)

Assumption Checking for the Mixed-Design ANOVA on the FITB Test

Before conducting the mixed-design ANOVA test, statistical assumptions associated with it were checked on the results of the FITB test. First, normality of the distribution of each variable was diagnosed via inspecting the value of skewness and kurtosis, the shape of histograms and boxplots, and results of the Shapiro-Wilk test (see Table 7 and 11; Figure 10 and 16). Eye-balling each histogram presented in Figure 7, all were negatively skewed, which was also supported by the descriptive data for skewness, which were less than the value of -1 (see Table #). The kurtosis values were much less than 3, which indicated relatively flat distributions except the value of the delayed posttest by the control group. In addition to these interpretations of the descriptive data, the results of the Shapiro-Wilk test for the EGE group and the control group also reached the significant value (see Table 7). Therefore, it has to be concluded that the assumption of normal distribution was violated in most of the dataset.

As for the variance of each variable, relatively different sizes of the boxplots implied that the data variance across the groups were not homogenous. However, results of the Levene's test did not reach significance, $F(8, 90) = 0.79$, $p = .61$, meaning that each variance of the groups was not significantly different from each other.

The distribution of the residuals was also diagnosed based on the model created for the mixed-design ANOVA test. Since data points of the residual did not look aligned along the straight diagonal line in Figure 11, the histogram of the model residuals was also examined (see Figure 12). Based on the visual inspection of these figures, it can be concluded that the distribution of the model residuals was not normal, which was confirmed by the results of the Shaprio-Wilk test on the residuals, $W = 0.88$, $p < .001$. As for the variance of the residuals, Figure 13 illustrates a scatterplot of the residuals and fitted values extracted from the model. As

the plot shows, they were not scattered below and above around the line of zero, thus indicating that the variance of the model residuals was not equal.

Since the same person's data were examined repeatedly in the mixed-design ANOVA, the equality of variances among a single participant's data was also inspected using Mauchly's sphericity test. The results did not reach the significant *p*-value, $W = 0.85$, $p = .10$, meaning that the variance of differences was relatively equal. Therefore, the assumption of sphericity was not violated.

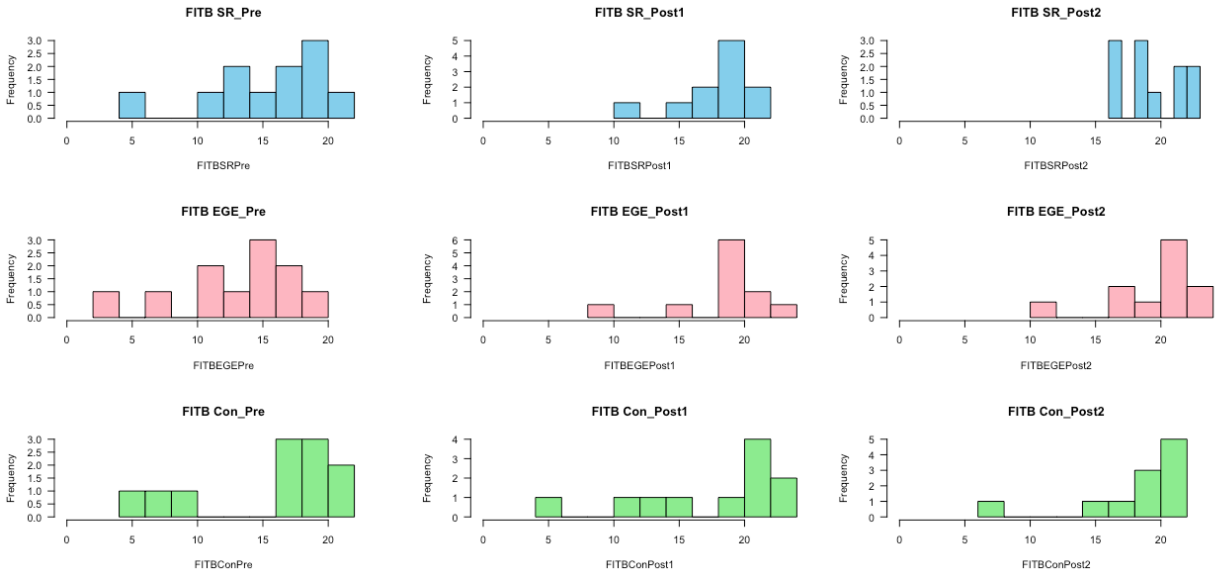


Figure 10. Histograms for the distribution of the FITB test scores

Table 7. *Results of Shapiro-Wilk test on the FITB test*

Instruction	Test	*W*	*n*	*p-value*
SR	Pretest	0.90	11	.18
	Posttest 1	0.86	11	.06
	Posttest 2	0.90	11	.20
EGE	Pretest	0.92	11	.31
	Posttest 1	0.74	11	.001
	Posttest 2	0.74	11	.002
Control	Pretest	0.82	11	.02
	Posttest 1	0.81	11	.01
	Posttest 2	0.72	11	.001

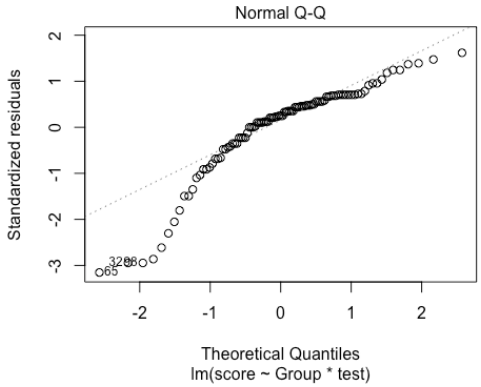


Figure 11. A normal qqplot for the residuals (FITB)

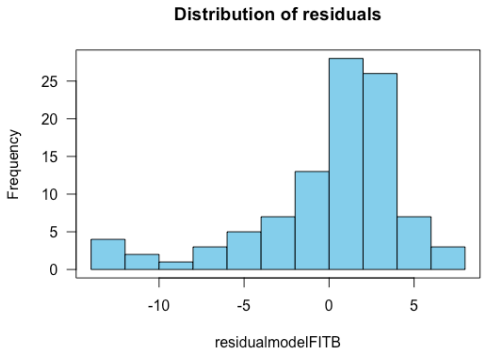


Figure 12. Histogram for the residuals (FITB)

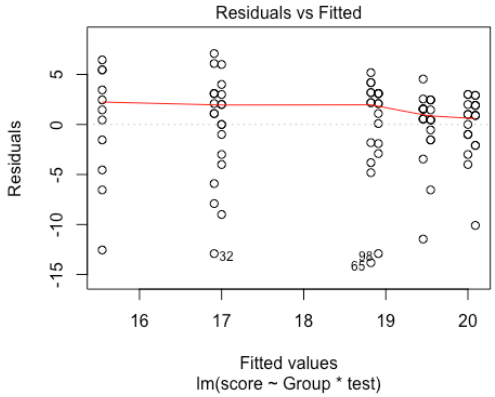


Figure 13. Residuals vs Fitted plots for the residuals (FITB)

CHAPTER 4

RESULTS

This chapter presents the results of the instructional treatments with the three different conditions (SR, EGE, and Comparison Group [CG]). To measure the effects of the instructional treatments, two different tests were administered: the Oral Elicited Imitation (OEI) test for measuring implicit knowledge (Ellis, 2005) and the Fill-in-the-Blank (FITB) test for measuring learners' explicit knowledge. Results from both OEI and FITB test will be presented in the following sections. For each testing measure, (1) descriptive summary and their visual counterparts are presented; (2) results of mixed-design ANOVA and of follow-up analyses are shown to compare group differences across the three testing sessions; (3) to address Research Question 2, the descriptive results of both tests on each target-linguistic (TL) form (i.e., RC[SU], RC[DO], and RC[OPREP]) are also presented with 95% confidence intervals (CIs) that compared the effects of instruction on each TL forms. Discussions of the findings presented in this chapter is taken up in the following chapter.

Results of the Oral Elicited Imitation (OEI) Test

Table 8 summarizes the descriptive results of each group's OEI test-performances. Visual counterparts of the descriptive results are presented in Figure 14 as boxplots. Each group's gain scores are also summarized in Table 9 and Figure 15.

As shown in Table 8 and Figure 14, the three groups all improved their OEI test-performances after the instructional treatments. As indicated by the mean scores and the gain scores, the SR group showed the highest increase at the immediate posttest, followed by the EGE group and the comparison group. However, at the delayed posttest, the EGE group caught up

with the SR group. Even the comparison group, which did not receive any form-focused instruction in the treatment, also showed a steady increase from the pretest to the delayed posttest. Due to the relatively large range of the 95% CIs, which may have been caused by the small sample size of each group, all of the 95% CIs among the groups and across the three testing phases overlapped one another. As indicated by Figure 14 that the increase from the pretest to the immediate posttest by the SR group looked somewhat different from the other two groups, it may be valuable to closely examine whether there is an interaction effect between Group and Time by conducting a mixed-design ANOVA test.

Table 8. *Descriptive statistics for the results of the OEI test*

Group	Test	*M*	*Mdn*	*Min*	*Max*	*SD*	*Skew*	*Kurtosis*	*95% CIs*
SR	Pretest	7.55	8	1	18	5.84	0.28	-1.53	[4.10, 11.00]
	Posttest1	12.00	13	3	21	5.74	-0.21	-1.39	[8.61, 15.39]
	Posttest2	13.27	14	6	24	5.78	0.18	-1.21	[9.86, 16.68]
EGE	Pretest	7.50	8	1	15	5.31	-0.08	-1.74	[4.21, 10.79]
	Posttest1	10.90	14	1	17	6.49	-0.43	-1.76	[6.88, 14.92]
	Posttest2	13.60	16	4	20	6.00	-0.48	-1.57	[9.88, 17.32]
CG	Pretest	8.18	9	0	20	6.42	0.38	-1.15	[4.40, 11.96]
	Posttest1	10.27	9	0	19	6.25	-0.02	-1.56	[6.59, 13.95]
	Posttest2	12.45	13	0	20	5.89	-0.65	-0.67	[8.96, 15.94]

Table 9. *Gain scores on the OEI test*

Group	Test	*M*	*Mdn*	*Min*	*Max*	*SD*	*Skew*	*Kurtosis*	*95% CIs*
SR-Gain	Posttest1	4.45	5	1	12	3.05	1.07	0.78	[2.65, 6.25]
	Posttest2	5.72	5	1	12	3.29	0.64	-0.74	[3.78, 7.66]
EGE-Gain	Posttest1	3.40	3.5	0	7	2.46	-0.01	-1.54	[1.87, 4.93]
	Posttest2	6.10	6.5	3	10	2.60	0.25	-1.5	[4.49, 7.71]
CG-Gain	Posttest1	2.09	3	-2	5	2.59	-0.19	-1.7	[0.56, 3.62]
	Posttest2	4.27	5	-2	9	3.32	-0.43	-1.07	[2.31, 6.23]

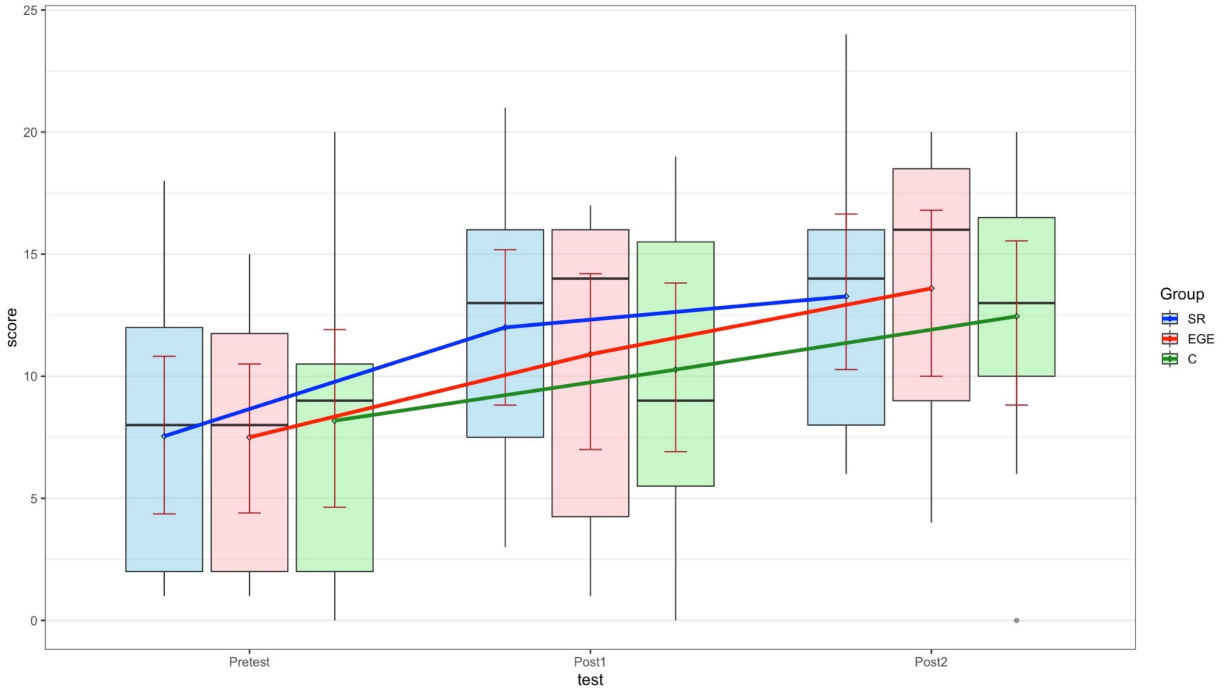


Figure 14. Boxplots for the three groups' test-performances on the OEI test

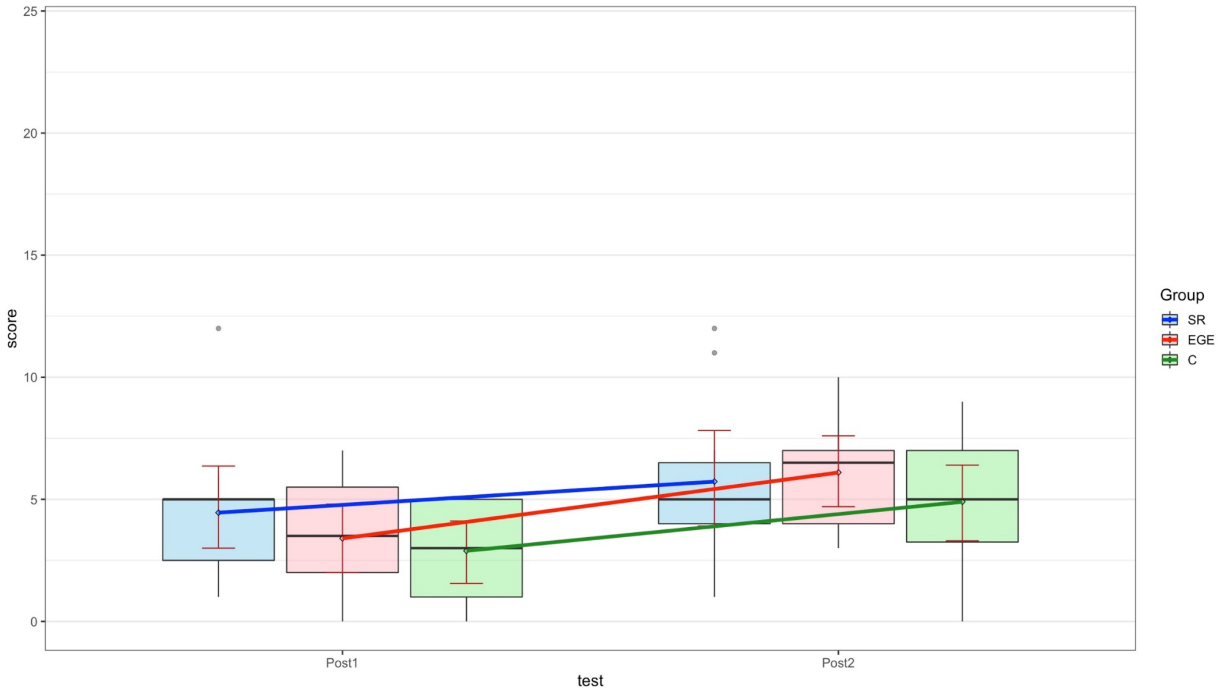


Figure 15. Boxplots for the gain scores on the OEI test

To examine the group differences across the three testing sessions on the learners' implicit knowledge development, a mixed-design ANOVA was conducted with Group (SR, EGE, and CG) as a between-subjects factor and Time (Pretest, Immediate posttest, and Delayed posttest) as a within-subjects factor. It has to be noted here that the sample sizes were very small, and thus the results of the parametric test need to be carefully interpreted in relation to the results of the assumption checking, the descriptive results, and its visual representations.

Before conducting the mixed design ANOVA, a one-way ANOVA was conducted to examine the group-differences at the point of pretest (see Table 10). As shown in Table 10, no statistically significant differences were found on the pretest-scores among the three instructional groups, $F(2, 29) = 0.05$, $p = .96$, $\eta^2 = 0.003$.

Table 10. *Results of one-way ANOVA on the pretest scores (OEI)*

Group	*N*	*M (SD)*	*F-statistic*	*p-value*	η^2	*df*
SR	11	7.55 (0.28)	0.05	.95	0.003	Between groups = 2
EGE	10	7.50 (5.31)				Within groups = 29
CG	11	8.18 (6.42)				

In the previous assumption checking process (see the results of assumption checking in the previous Method Chapter), only the assumptions of equal variance and sphericity were satisfied. It has to be admitted that the power of the statistical test was weak in the analysis due to the small number of participants (Larson-Hall, 2016). However, to explore the general tendency of the results and to supplement the interpretation of the descriptive results, a mixed-design ANOVA was performed to compare the group means across the three testing phases. The results of the ANOVA test were interpreted and discussed carefully along with the descriptive statistics and the visual representations of the results in the next chapter.

Contrary to the original hypothesis, results of the mixed-design ANOVA revealed that a main effect for Time (Pretest, Immediate posttest, and Delayed posttest), $F(2, 58) = 566.48$, $p < .001$, $\eta^2 = 0.13$, whereas the effect for Group (SR, EGE, and CG), $F(2, 29) = 0.03$, $p = .97$, $\eta^2 = 0.002$, and the interaction between Group and Time, $F(4, 58) = 1.25$, $p = .29$, $\eta^2 = 0.007$, did not show any significant results. Based on these results, all the instructional group may have improved their performances similarly and no statistical differences were found depending on the different instructional conditions.

Results of the Fill-in-the-Blank (FITB) Test

Descriptive statistics from the test-scores by each instructional group are summarized in Table 11. The gain scores of each group are presented in Table 12. The visual counterparts of the results are shown as boxplots (see Figure 16 and 17).

Similar to the results of the OEI test, all of the three instructional groups improved their test-performances in the posttest phases, and their gain scores were retained at the delayed posttest administered two weeks after the final instructional treatment. As shown by Table 11 and 12 and Figure 16 and 17, the provision of explicit grammar explanation seemed to have the biggest impact on the learners' test-performances among the three instructional groups, seconded by the SR task. Both total and gain scores by the comparison group were the lowest among the three groups.

Although the mean scores for the EGE group and the SR group were higher than the comparison group, the 95% CIs of the three groups greatly overlapped one another (see Table 11). Comparing 95% CI values within each group, they overlapped one another except the difference between the pretest and the delayed posttest for the EGE group. Thus, it can be

predicted that there may be a statistically significant difference between these two test-results. Likewise, the results of the gain scores indicated that there were substantial overlaps (see Table 12). However, the ranges of the 95% CIs on the immediate posttest by the EGE group and the comparison group did not overlap, indicating a significant difference between these two groups' gain scores. Although the within-group and between-group differences can be predicted by comparing the 95% CIs, the discussions need to be made carefully due to the small sample size of each instructional group, which was likely to have contributed to the ranges of the CIs being large.

Table 11. *Descriptive statistics for the results of the FITB test*

Instruction	Test	*M*	*Mdn*	*Min*	*Max*	*SD*	*Skew*	*Kurtosis*	*95% CIs*
SR	Pretest	15.73	17	5	21	4.76	-0.82	-0.31	[12.93, 18.53]
	Posttest1	18.45	19	11	22	3.01	-1.20	0.70	[16.67, 20.23]
	Posttest2	19.64	19	16	23	2.62	-0.06	-1.61	[18.09, 21.19]
EGE	Pretest	13.64	16	3	20	5.12	-0.71	-0.72	[10.62, 16.66]
	Posttest1	18.91	20	8	24	4.18	-1.49	1.49	[16.44, 21.38]
	Posttest2	20.00	21	10	23	3.74	-1.62	1.75	[17.79, 22.21]
CG	Pretest	16.00	18	4	22	6.29	-0.80	-1.09	[12.28, 19.72]
	Posttest1	17.91	21	4	23	6.20	-1.00	-0.39	[14.24, 21.58]
	Posttest2	18.64	20	6	22	4.65	-1.71	2.04	[15.90, 21.38]

Table 12. *Gain scores on the FITB test*

Instruction	Test	*M*	*Mdn*	*Min*	*Max*	*SD*	*Skew*	*Kurtosis*	*95% CIs*
SR-Gain	Posttest1	2.73	2	-1	7	2.76	0.47	-1.36	[1.10, 4.36]
	Posttest2	3.90	3	0	12	3.53	0.86	-0.12	[1.80, 6.00]
EGE-Gain	Posttest1	5.27	5	1	12	3.17	0.67	-0.53	[3.41, 7.13]
	Posttest2	6.36	7	0	13	3.41	-0.06	-0.32	[4.34, 8.38]
CG-Gain	Posttest1	1.91	1	-1	5	2.02	0.18	-1.65	[0.71, 3.11]
	Posttest2	2.63	2	0	9	3.20	0.93	-0.69	[0.73, 4.53]

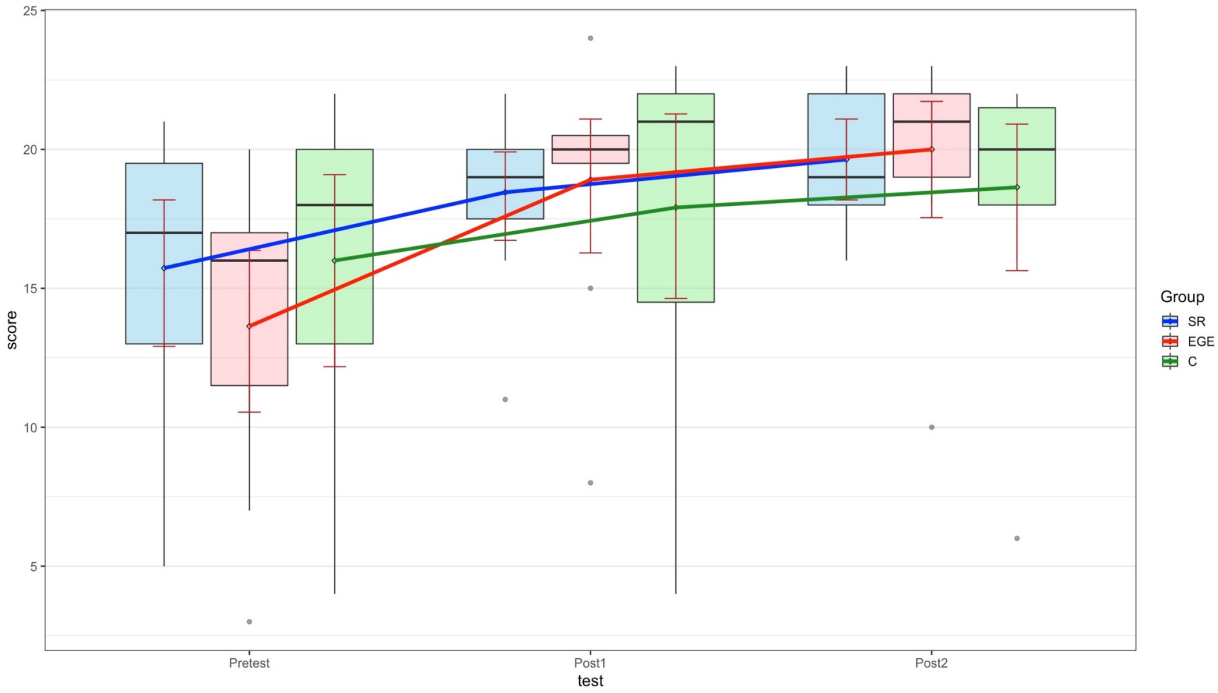


Figure 16. Boxplots for the three groups' test-performances on the FITB test

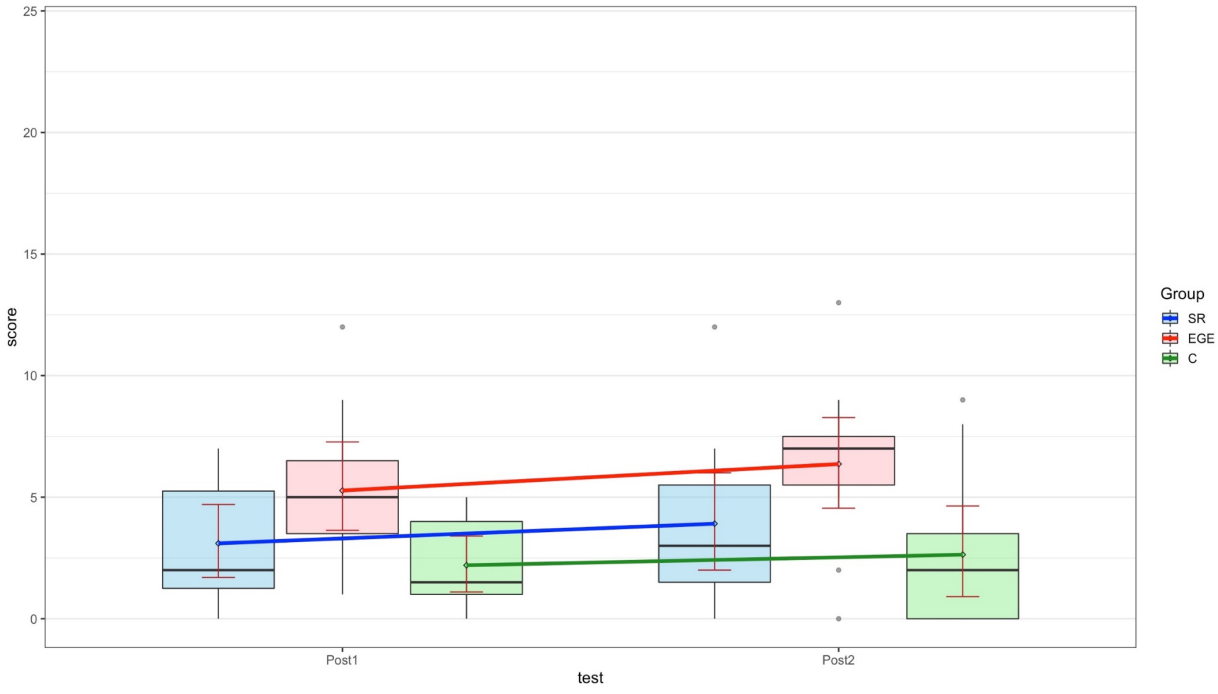


Figure 17. Boxplots for the gain scores on the FITB test

To understand the group differences across the three testing sessions, a mixed-design ANOVA was performed with Group (SR, EGE, and CG) as a between-subjects factor and Time (Pretest, Immediate posttest, and Delayed posttest) as a within-subjects factor. First, to examine whether the three groups were equivalent at the pretest, a one-way ANOVA was performed on their pretest-scores. Since the normality of the dependent variables and the homogeneity of variance were discussed above, only the results of Levene's test is presented here, $F(2, 30) = 0.10$, $p = .91$. As shown in Table 13, the results of the one-way ANOVA did not show any statistically significant result, $F(2, 30) = 0.62$, $p = .54$, $\eta^2 = 0.04$, indicating that the groups were not statistically different from each other at the pretest.

Table 13. *Results of one-way ANOVA on the pretest scores (FITB)*

Group	*N*	*M (SD)*	*F-statistic*	*p-value*	η^2	*df*
SR	11	15.73 (4.76)	0.62	.54	0.04	Between groups = 2
EGE	11	13.64 (5.12)				Within groups = 30
CG	11	16.00 (6.29)				

As shown in the previous assumption checking (see the results of assumptions checking in the previous Method Chapter), some assumptions for conducting a mixed-design ANOVA were not met. As discussed in the analysis of the OEI test, a mixed-design ANOVA was still conducted on the results of the FITB test for the purpose of exploring the general tendencies and supplementing the descriptive data.

Results of the mixed-design ANOVA revealed that there was a significant main effect for Time, $F(2, 60) = 40.16$, $p < .001$, $\eta^2 = 0.15$, and an interaction effect between Group and Time, $F(2, 60) = 2.96$, $p = .03$, $\eta^2 = 0.02$, but there was no significant effect for Group, $F(2, 30) = 0.03$, $p = .97$, $\eta^2 = 0.002$. To specify where exactly the differences were across the three testing phases and among the groups, first, the line-graph presented in Figure 16 was examined, and one-way

ANOVA with repeated measure was then performed on Time as a within factor, followed by post-hoc pair-wise *t*-tests. As Figure 16 shows, the test-performances by all the three groups increased from pretest to immediate posttest and the scores further increased slightly at the delayed posttest. Based on a visual observation, it can be predicted that the differences in Time may lie between the pretest and the immediate posttest, and between the pretest and the delayed posttest, but not between the immediate and the delayed posttest. Results of the follow-up one-way ANOVA, $F(2, 96) = 7.98$, $p < .001$, $\eta^2 = 0.14$, and the pair-wise post hoc test confirmed the visual observation (see Table 14).

Table 14. *Results of post hoc pair-wise t-test, effect sizes, and 95% CIs for Time*

Time	*Bonferroni (p-value)*	*Cohen's d*	*95% CIs*
Pretest-Posttest1	< .001	-0.67	[-1.17, -0.16]
Pretest-Posttest2	< .001	-0.93	[-1.45, -0.42]
Posttest1-Posttest2	.07	-0.24	[-0.74, 0.25]

Next, a series of one-way ANOVAs were conducted on the gain scores from the pretest to the immediate posttest and from the pretest to the delayed posttest (see Table 12 for the descriptive statistics on the gain scores). The results of the follow-up one-way ANOVAs are presented in Table 15 and 17, both of which showed significant differences among the groups. To specify where the differences were, Bonferroni post-hoc pairwise comparisons were conducted. On the gain scores in the immediate posttest, the difference between the EGE group and the comparison group showed a significant value with a large effect size (see Table 16). As pointed out based on the descriptive results, the 95% CIs for the group-differences also supports the results of the previous pairwise post-hoc test. The range of higher and lower CI values for the differences between the EGE group and the comparison group did not cross zero, whereas the other CI values for the other pairs (i.e., SR-EGE and SR-CG) crossed zero (see Table 16). The

similar tendency was found on the results of the post-hoc pairwise comparisons on the gains scores of the delayed posttest (see Table 18). The group difference between the EGE group and the comparison group was the only pair that showed a significant value, and the 95% CIs did not cross zero. In sum, the results of the mixed-design ANOVA test and the follow-up comparisons revealed statistically significant differences and the large effect sizes on the gain scores between the EGE group and the exposure only comparison group.

Table 15. *Results of the follow-up one-way ANOVA on the gain scores (Pretest-Posttest1)*

Group	*N*	*M (SD)*	*F-statistic*	*p-value*	η^2	*df*
SR	11	2.73 (2.76)	4.674	.01	0.24	Between groups = 2
EGE	11	5.27 (3.17)				Within groups = 30
CG	11	1.91 (2.02)				

Table 16. *Post-hoc pairwise comparisons on the gain-scores (Pretest-Posttest1)*

Group	*M Difference*	*Lower 95% CIs*	*Higher 95% CIs*	*p-value*	*Cohen's d*
SR - EGE	-2.54	-1.79	0.07	.10	-0.86
SR - CG	0.82	-0.56	1.23	1.00	-0.34
EGE - CG	3.36	0.29	2.24	.02	1.27

Table 17. *Results of the follow-up one-way ANOVA on the gain scores (Pretest-Posttest2)*

Group	*N*	*M (SD)*	*F-statistic*	*p-value*	η^2	*df*
SR	11	3.90 (3.53)	3.443	.045	0.19	Between groups = 2
EGE	11	6.36 (3.41)				Within groups = 30
CG	11	2.63 (3.20)				

Table 18. *Post-hoc pairwise comparisons on the gain-scores (Pretest-Posttest2)*

Group	*M Difference*	*Lower 95% CIs*	*Higher 95% CIs*	*p-value*	*Cohen's d*
SR - EGE	-2.46	-1.62	0.21	.30	-0.71
SR - Control	1.27	-0.52	1.27	1.00	0.38
EGE - CG	3.73	0.17	2.08	.045	1.13

Descriptive Statistics of Both (FITB and OEI) Tests on Each Relative Clause (RC) Type

In order to address Research Question 2 which aimed to examine whether the effects of instruction differ depending on the types of RCs (SU, DO, and OPREP), results of both tests on each RC type were also analyzed. As discussed throughout this chapter, the number of the participants in this study was very small. In addition, multiple independent variables (i.e., RC Type, Group, and Time) were included in the analyses to address Research Question 2. Thus, the analyses here focused on the detailed examination of the tendencies indicated by the descriptive results, the visuals and the 95% CIs, rather than performing parametric statistical tests. Each value for the 95% CIs was also compared between groups, RC types, and testing-phases. First, results of both OEI and FITB tests on each RC type and results of visual observations of these test-results are presented. Then, gain scores on each RC type were examined to supplement the analyses.

Results of the OEI Test on Each RC Type

Results of the OEI test on each RC type also showed higher scores at the posttest sessions by the two experimental groups (SR and EGE) than by the comparison group (see Table 19 and Figure 18). As shown in Table 19, all the 95% CIs overlapped one other between groups, RC types, and across all the testing-sessions. Even though the 95% CIs did not show any significant differences on the learners' test-performances depending on the RC types, their performances on the most marked OPREP looked lower than their performances on the less marked RC types (SU and DO).

Concerning the gains scores on each RC type (see Table 20 and Figure 19), both experimental groups (SR and EGE) seemed to be equally more advantageous on their

performances on the least marked form (i.e., SU) than the comparison group. However, the differences between the experimental groups and the comparison group was not observed on more marked forms (DO, OPREP).

Results of the FITB Test on Each RC Type

Table 21 and Figure 18 show the results of the FITB test on each RC type (SU, DO, and OPREP). As indicated by the mean scores, both experimental groups (SR and EGE) generally showed higher performances on their posttests compared to the comparison group. However, all the 95% CIs overlapped each other between groups across the pretest to the delayed posttest. Even though all the 95% CIs overlapped each other between groups, the experimental groups (SR and EGE) improved their performances substantially from the pretest to the delayed post-test to the extent that their 95% CIs showed no overlap or very slight overlap, especially on SU and DO (see Table 21). Comparing the test-results on each RC type, the test-performances by all three groups on SU and DO were higher than their performances on OPREP, which was a more marked form than the other two forms. The SR group's delayed posttest performances showed a significant difference between SU and OPREP based on the comparison of their CI values.

The results of the gains scores on the FITB test also showed similar tendencies observed in Table 20: both experimental groups (SR and EGE) gained higher scores than the comparison group based on the descriptive statistics (see Table 20 and Figure 19). One interesting tendency was that, especially on the most marked type of RC (i.e., OPREP), providing explicit grammar explanation seemed more advantageous than the other instructional conditions as indicated by the EGE group's gain scores on OPREP, which was quite contrary to the results of the OEI test. The 95% CIs of the gain scores between the EGE group and the comparison group in the FITB

test did not overlap each other, indicating the significant difference. Of course, a firm conclusion cannot be made based on these two observations, due to the small sample size and lack of parametric tests. However, these results may imply some differential effects of the two instructional conditions (SR and EGE) on different types of L2 knowledge (i.e., explicit and implicit knowledge). The interpretations of these analyses are discussed in the following discussion chapter.

Table 19. *Results on the OEI test on each of the three RC types*

		Pretest				Posttest1				Posttest2			
RC Type	*Group*	*M*	*Mdn*	*SD*	*95% CIs*	*M*	*Mdn*	*SD*	*95% CIs*	*M*	*Mdn*	*SD*	*95% CIs*
SU	SR	2.63	2	2.16	[1.36, 3.90]	4.27	4	2.45	[1.36, 5.72]	4.82	5	1.99	[3.64, 6.00]
	EGE	2.60	3	1.43	[1.72, 5.72]	3.90	4	2.18	[2.54, 5.25]	5.00	5.5	2.26	[3.61, 6.39]
	Con	3.45	4	2.54	[1.94, 4.96]	3.36	3	2.69	[1.77, 4.95]	4.63	5	2.16	[3.36, 5.90]
DO	SR	3.45	5	2.21	[2.14, 4.76]	4.91	6	2.02	[3.71, 6.11]	5.18	7	2.48	[3.71, 6.65]
	EGE	2.70	3	2.11	[1.39, 4.01]	4.20	5.5	2.94	[2.38, 6.02]	5.10	5.5	2.13	[3.79, 6.41]
	Con	3.09	2	2.59	[1.56, 4.62]	4.36	4	2.11	[3.11, 5.61]	4.73	5	0.75	[3.26, 6.20]
OPREP	SR	1.45	0	1.81	[0.37, 2.53]	2.82	3	1.99	[1.64, 4.00]	3.27	4	2.15	[2.00, 4.54]
	EGE	2.20	2	2.04	[0.93, 3.47]	2.80	3	2.04	[1.53, 4.07]	3.50	4	1.84	[2.36, 4.64]
	Con	1.64	1	1.86	[0.54, 2.74]	2.54	2	2.16	[1.27, 3.81]	3.09	3	1.76	[2.05, 4.13]

Possible maximum = 8

Table 20. *Results on the Gain Scores on the OEI test on each of the three RC types*

		Pretest-Posttest1				Pretest-Posttest2			
RC Type	*Group*	*M*	*Mdn*	*SD*	*95% CIs*	*M*	*Mdn*	*SD*	*95% CIs*
SU	SR	1.64	1	2.16	[0.37, 2.91]	2.18	2	1.54	[1.28, 3.08]
	EGE	1.30	1	1.49	[0.38, 2.22]	2.40	2.5	1.35	[1.56, 3.24]
	Con	-0.09	0	1.22	[-0.82, 064]	1.18	1	1.68	[0.49, 1.87]
DO	SR	1.45	1	1.04	[0.84, 2.06]	1.73	2	1.74	[0.71, 2.75]
	EGE	1.50	1.5	1.35	[0.66, 2.34]	2.40	2	1.17	[1.67, 3.13]
	Con	1.27	2	1.49	[0.39, 2.15]	1.64	1	1.50	[0.76, 2.52]
OPREP	SR	1.36	2	1.21	[0.65, 2.07]	1.82	1	1.40	[1.00, 2.64]
	EGE	0.60	1	0.84	[0.07, 1.13]	1.30	1	1.16	[0.57, 2.03]
	Con	0.91	1	1.51	[0.01, 1.81]	1.45	2	1.44	[0.61, 2.29]

Possible maximum = 8

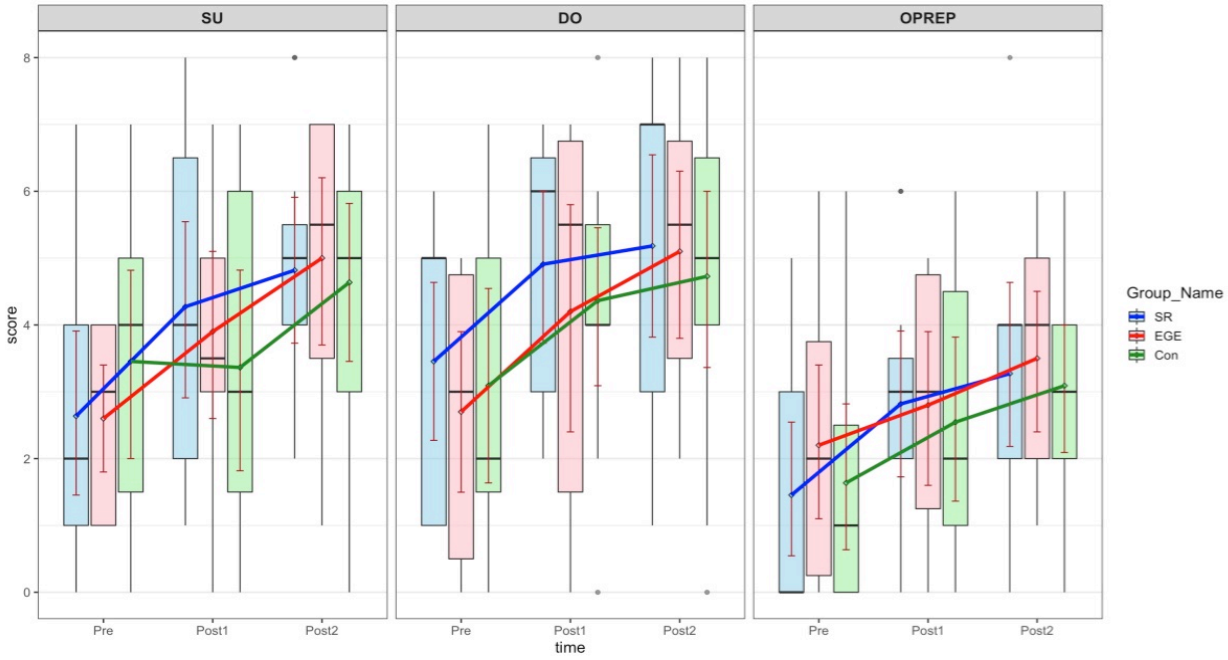


Figure 18. Boxplots for the results on the OEI test on each of the three RC type

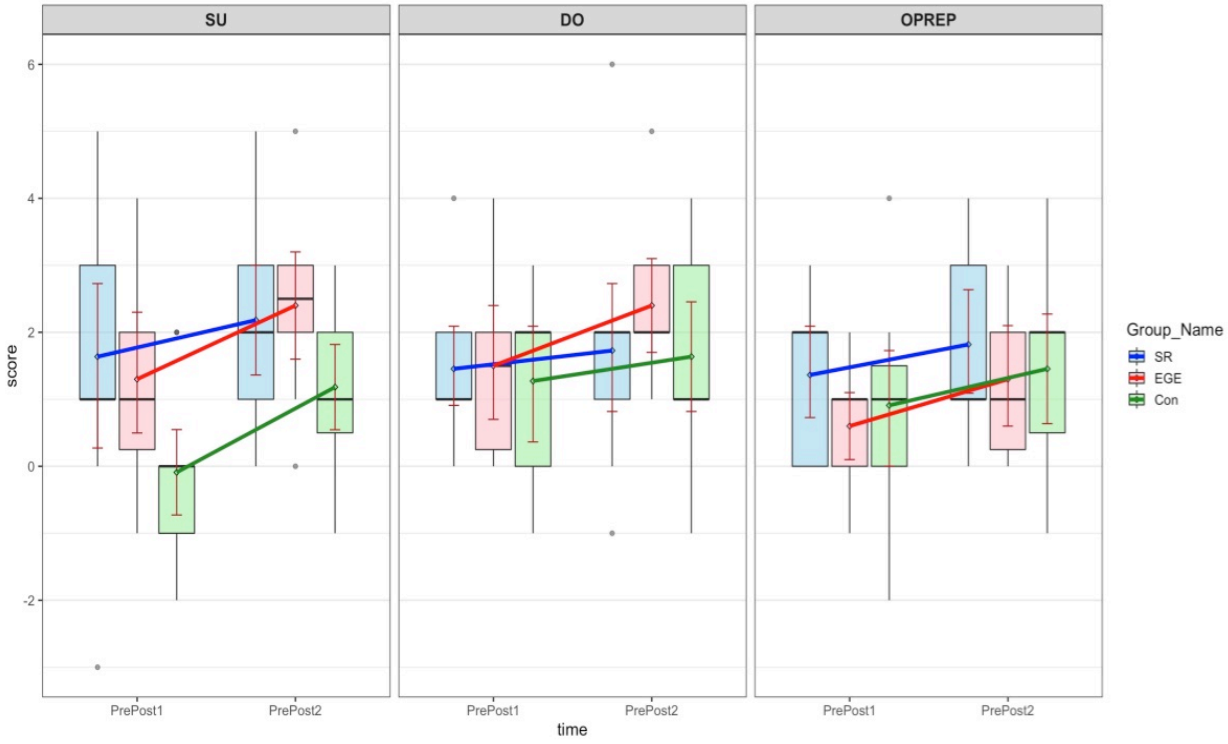


Figure 19. Boxplots for the gain scores on the FITB test on each of the three RC types

Table 21. *Results on the FITB test on each of the three RC types*

		Pretest				Posttest1				Posttest2			
RC Type	*Group*	*M*	*Mdn*	*SD*	*95% CIs*	*M*	*Mdn*	*SD*	*95% CIs*	*M*	*Mdn*	*SD*	*95% CIs*
SU	SR	5.66	*7*	2.20	[4.37, 6.95]	6.45	7	1.21	[5.72, 7.15]	7.27	7	0.79	[6.80, 7.74]
	EGE	5.45	6	1.75	[4.41, 6.49]	7.18	8	1.53	[6.27, 8.08]	7.45	8	0.82	[6.96, 7.94]
	Con	5.45	6	2.46	[4.00, 6.90]	6.45	8	2.16	[5.18, 7.72]	7.00	7	1.34	[6.22, 7.78]
DO	SR	4.82	5	1.89	[3.70, 5.94]	6.27	7	1.49	[5.39, 7.15]	6.91	7	1.04	[6.30, 7.52]
	EGE	4.09	4	1.97	[2.93, 5.25]	5.91	6	1.45	[5.05, 6.77]	6.55	7	2.07	[5.33, 7.77]
	Con	5.27	6	2.65	[3.70, 6.84]	6.09	7	2.51	[4.60, 7.58]	6.00	7	2.10	[4.77, 7.23]
OPREP	SR	5.18	5	1.66	[4.20, 6.16]	5.73	6	1.01	[5.14, 6.32]	5.64	6	1.43	[4.80, 6.48]
	EGE	4.09	4	2.34	[2.70, 5.48]	5.91	6	1.70	[4.91, 6.91]	6.00	6	1.18	[5.29, 6.71]
	Con	5.27	6	2.15	[4.00, 6.54]	5.45	6	1.63	[4.49, 6.41]	5.63	6	1.68	[4.63, 6.63]

Possible maximum = 8

Table 22. *Results on the gain scores on the FITB test on each of the three RC types*

		Pretest-Posttest1				Pretest-Posttest2			
RC Type	*Group*	*M*	*Mdn*	*SD*	*95% CIs*	*M*	*Mdn*	*SD*	*95% CIs*
SU	SR	0.73	0	1.19	[0.02, 1.44]	1.55	1	1.92	[0.41, 2.69]
	EGE	1.73	2	1.27	[0.99, 2.47]	2.00	2	1.18	[1.29, 2.71]
	Con	1.00	1	1.34	[0.22, 1.78]	1.55	1	2.21	[0.24, 2.86]
DO	SR	1.45	2	1.51	[0.57, 2.33]	2.09	2	1.64	[1.13, 3.05]
	EGE	1.82	2	1.25	[1.08, 2.56]	2.45	2	1.86	[1.35, 3.55]
	Con	0.82	1	1.47	[-0.04, 1.68]	0.73	1	1.35	[-0.07, 1.53]
OPREP	SR	0.55	1	1.57	[-0.37, 1.47]	0.45	0	1.51	[-0.43, 1.33]
	EGE	1.82	2	2.04	[0.60, 3.04]	1.91	3	2.07	[0.69, 3.13]
	Con	0.18	0	0.87	[-0.33, 0.69]	0.36	0	0.92	[-0.19, 0.91]

Possible maximum = 8

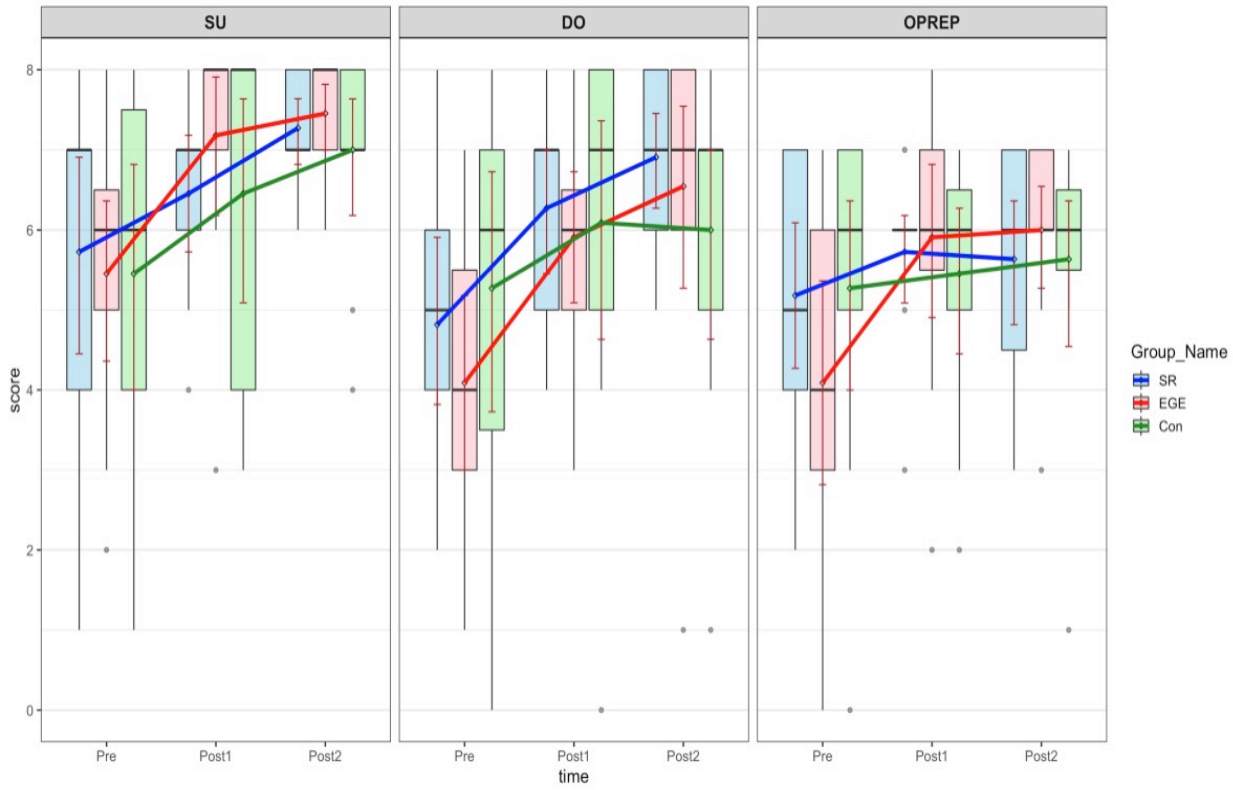


Figure 20. Boxplots for the results on the FITB test on each of the three RC types

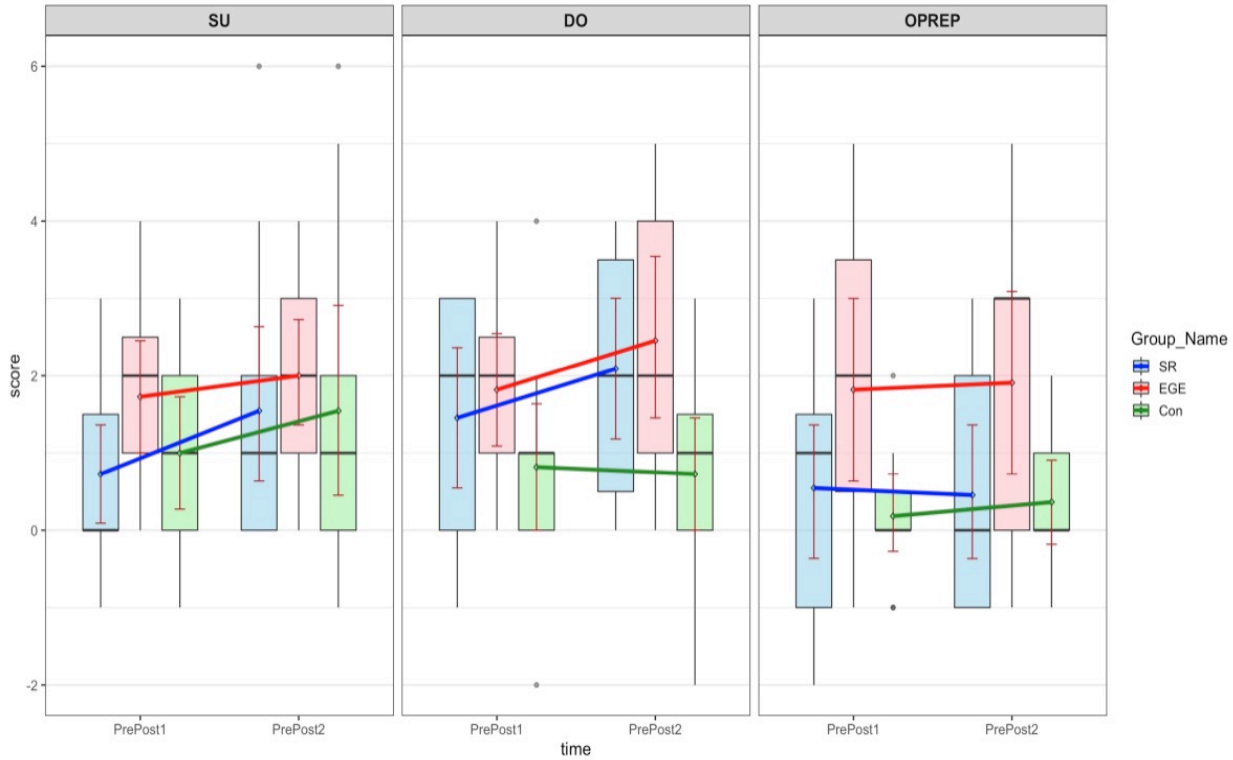


Figure 21. Boxplots for the gain scores on the FITB test on each of the three RC types

CHAPTER 5

DISCUSSION

This chapter provides detailed discussions on the results reported in the previous chapter. In addition to main discussions with reference to each research question, further analyses and discussions are also provided to closely examine the effects of the story-retelling (SR) task and the provision of explicit grammar explanation on learners' implicit and explicit knowledge development.

The primary motivation for this study was to investigate whether output-based focus on form that was systematically constructed based on psycholinguistic rationales and previous ISLA research findings contributes to the development of L2 learners' implicit knowledge, a type of linguistic knowledge that can be used quickly in spontaneous communication (DeKeyser, 2003). To address a long-standing measurement issue in previous effect-of-L2-instruction research (Ellis, 2005; Doughty, 2003; Norris & Ortega, 2000), this study employed two different tests (i.e., the oral elicited imitation [OEI] test and the fill-in-the-blank [FITB] test) to measure relative impacts of the instruction on L2 learners' implicit and explicit knowledge. Furthermore, to explore possible strengths and limitations of two different types of form-focused instruction, the effects of the SR task and the provision of explicit grammar explanation were compared with each other and with the results by the exposure only comparison condition. In relation to differential effects of two types of form-focused instruction, how the complexity of the target grammatical forms interacts with the effects of instruction was also examined by targeting the three types of RCs: Subject-type (SU), Direct-object-type (DO), and Object-of-preposition-type

(OPREP). The following discussions will address the two research questions based on the interpretation of the results of both OEI and FITB tests.

Effects of SR and Explicit Explanation on Learners' Implicit Knowledge Development

Contrary to the researcher's original expectations, the SR group failed to outperform either of the other two instructional groups (the EGE group and the comparison groups) across the three testing sessions of the OEI test. As shown in the results of the OEI test, the results of the mixed-design ANOVA only showed a significant main effect for Time but not for Group and Interaction between Time and Group. In other words, all the instructional groups improved their test-performances from the pretest to the posttests regardless of their instructional conditions. Therefore, the SR treatment did not make unique contribution to the development of the learners' implicit knowledge compared to the other instructional groups (the EGE group and the comparison group). Admittedly, it is difficult to make strong arguments with the quantitative evidence due to the small sample size and low statistical power for the inferential statistics. However, it is very important to carefully discuss why all the groups showed similar improvements in the posttest sessions even though they received different types of instruction. Especially, the comparison group did not even receive any form-focused instruction during the instructional treatments but significantly improved in the posttest sessions.

Several possible explanations can be considered to account for the unexpected gains made by the comparison group. The first explanation can be attributed to the instructional condition for the comparison group. Strictly speaking, the comparison group in this study was not a test-only control group. Rather, they were also exposed to the instances of RCs in the reading-texts multiple times (54 times in total) throughout the entire three treatment sessions. In

particular, the target linguistic forms were not completely new to the participants as indicated by their pretest-scores both on the OEI and FITB tests. In other words, the learners already possessed partial knowledge of the RCs before the instructional treatments began. Therefore, even though the group did not receive any form-focused instruction, it is possible that the learners benefitted from the mere exposure to the input that contained the instances of the target linguistic forms and that the effects appeared gradually over two weeks after the treatment sessions (Doughty, 1990; Muranoi, 2000).

Another possible, but the most plausible explanation may be related to a test-practice effect caused by taking the OEI test multiple times throughout the whole testing sessions. Ellis (2009) pointed out this potential limitation as one of the caution points of the OEI test. Previous studies that used the OEI test to measure the effects of form-focused instruction on L2 learners' implicit knowledge with a pretest and posttests design (e.g., Ellis, Loewen, & Erlam, 2006; Loewen, Erlam, & Ellis, 2009) reported some gains made by the control group in the delayed posttest, though their experimental groups outperformed the control group. Carefully re-examining the procedures of the OEI test, it turned out that the test itself may have functioned as a potential learning opportunity. In the OEI test used in the present study, the test consisted of the following four procedures: (1) listening to a stimulus statement; (2) comprehending the meaning of the statement to answer a following belief questionnaire; (3) answering the belief questionnaire by choosing '*True*', '*False'*, or '*Not sure*' for the learners; and (4) repeating the statement based on the meaning that the learners comprehended. Since the second and third procedures aimed at directing learners' focus on the meaning of the statement to satisfy one of the important criteria (i.e., focus on meaning) to tap into learners' implicit knowledge (see Chapter 2 for the detailed description of the key-factors to create more valid measuring-tools for

implicit and explicit knowledge; also see Ellis, 2015; Erlam, 2006), it can be assumed that the OEI test itself might have served as further practice that involved both meaningful input- and output-processing practices. Considering the fact that the target linguistic forms were not new to the students, it may be likely that the test-practice effect caused by the design of the OEI test influenced the comparison group's test-performances in the posttest sessions. If so, the improvement made by the comparison group may have been mediated by the combination of their pre-existing knowledge, meaning-focused input in the treatment sessions, and the OEI performances in the testing sessions. If these explanations were the case, it was also possible to assume that the EGE group's highest gains in the delayed posttest may be, to certain extent, attributed to the combination of test-practice effects, their pre-existing knowledge, meaning-focused input and explicit grammar explanation provided in the treatment sessions. Indeed, previous ISLA studies that investigated the effects of form-focused instruction reported beneficial effects of providing explicit grammar explanation not only on L2 learners' explicit knowledge but also on their implicit knowledge (e.g., Ellis, Loewen, & Erlam, 2006; Li, Ellis, & Zhu, 2015).

However, it is impossible to specify which explanation (or the combination of these two explanation) was the case based on the research design because the study did not include any test-only control group. In addition, it is also possible that the participants became familiar with taking the OEI test just by taking the test multiple times. On the basis of this, an important methodological direction regarding the use of the OEI test must be pointed out. That is, if an effect-of-instruction study employs the OEI test with the aim of measuring the impact of instruction on learners' implicit knowledge development, a test-only control group must be included to identify how much gains can be accounted for by instructional treatments. Although

previous test-validation research has examined the validity and the reliability of the OEI test as a measure of implicit knowledge (e.g., Ellis, 2005; Ellis & Loewen, 2007; Spada et al., 2015), its potential test-practice effect reported in this study was not so much emphasized. The findings of the study pointed out the importance of investigating whether (or how much) the OEI test has learning effects on the results of posttests, especially the result of a delayed posttest. This type of research can be done by comparing the potential effects of the following variables: (1) types of target-linguistic forms (e.g., a form that requires syntactic manipulations and a form that requires morphological processing); (2) proficiency levels of learners (e.g., learners who possess a certain amount of pre-existing knowledge of the target-linguistic form and learners who do not); and (3) the number of total test-items. Future research that explores these questions will provide further implications on the interpretation of the present study.

Effects of SR and Explicit Explanation on Learners' Explicit Knowledge Development

In contrast to the results of the OEI test, the mixed design ANOVA test on the results of the FITB test indicated a significant interaction effect between Group and Time. The post-hoc comparisons on the gain scores revealed statistically significant differences between the EGE group and the comparison group in both posttests. Particularly, it is interesting to note that the gains made by the EGE group in the FITB test were apparent right after the instructional sessions, whereas the EGE group's immediate gains made in the OEI test were not as much evident as the gains in the FITB test. From this finding, it can be claimed that providing explicit grammar explanation is a very effective instructional option to introduce form-focused instruction and has strong immediate impacts on the learning of RCs, especially when the goal of the instruction is to develop L2 learners' explicit knowledge. According to the results shown in

the present study, this is one of the most evident strengths of providing explicit grammar explanation compared to the other instructional conditions: the SR task and the exposure only comparison condition.

Although the results discussed above accorded with the initial hypothesis, what was not expected was that the highest gains achieved by the EGE groups at the delayed posttest not only in the FITB test but also in the OEI test. Based on the results, an important argument that could be drawn from the results was that explicit grammar explanation, which is often categorized as focus on forms (FonFs), may also have had positive effects on the acquisition of implicit knowledge and contributed to the highest gains in the delayed posttests. Therefore, the findings of the study support Ellis' (2015) claim, "instruction that includes an explicit explanation of the target feature is, on the whole, more likely to be effective than instruction that does not" (p. 264).

Effects of Different Complexity of the RC Types

The second research question was: Do the two types of instructions (SR and explicit grammar explanation) have differential effects on the development of two types of different knowledge (implicit and explicit knowledge) depending on the complexity of the target linguistic forms: RC (SU), RC (DO), and RC (OPREP)? Concerning this research question, the original hypothesis predicted that the SR task would improve the learners' OEI test-performances on all the three types of RCs at both immediate and delayed post-tests, while the EGE group would improve their OEI test-performances only on the simpler and less-marked RC types: RC (SU) and RC (DO), but not on RC (OPREP). In addition, these gains on the simpler forms in the OEI test would have been predicted to disappear by the delayed posttest. The original hypothesis set up for the research question was partially confirmed.

First of all, the hypothesis was developed based on the degree of complexity and processing difficulty predicted by Keenan and Cormrie's (1977) Noun Phrase Accessibility Hierarchy (NPAH) and by the results of Izumi's (2003b) study. As shown by the results of both OEI and FITB tests, the learners' test-performances on the marked OPREP were lower than the other less-marked RC types (SU and DO). Comparing the results of the OEI test and the FITB test, the influence by the complexity (or markedness) of the RC types was evident in the OEI test. The results of the learners' test-performances on each RC type in both tests supported the predictions made by the previous studies on the degree of difficulty and markedness (Doughty, 1990; Izumi, 2003b; Keenan and Cormrie, 1977); thereby, it can be claimed that the marked OPREP was more difficult than less marked SU and DO.

Depending on the degree of complexity and difficulty of the RC types, effects of instruction also appeared to be affected. In the FITB test, both SR and EGE groups showed statistically significant improvement from the pretest to the delayed posttest on the less marked RC types (SU and DO), whereas only the EGE group showed clear gains on more marked OPREP in the FITB test. As indicated by the significant gain-scores in the FITB test, providing explicit grammar instruction was found to be beneficial on the learning of all RC types regardless of the different complexity of the RC types, whereas the impact of explicit explanation on the OEI test-performances seemed to be limited compared to the gains shown in the FITB test.

CHAPTER 6

CONCLUSION

This chapter first restate the aims of the study and present a summary of the key findings and the significance of these findings. In addition, potential limitations of the study and directions and recommendations for future research are discussed subsequently.

Summary of the Key Findings of the Study

The primary objective of this study was to investigate whether output-based focus on form, termed Story-Retelling (SR), makes a unique contribution to the development of Japanese EFL learners' implicit knowledge of English relativization. To explore this main question, the impact of the SR task was compared with the provision of explicit grammar explanation and the exposure only comparison condition. The effects of these instructional conditions were measured through the OEI and FITB tests for measuring learners' implicit and explicit knowledge, respectively. In addition to the main research question, this study also explored how the complexity of the target linguistic forms influence the effects of form-focused instruction by targeting the three types of RCs: Subject-type (SU), Direct-object-type (DO), and Object-of-preposition-type (OPREP).

Based on findings of previous ISLA studies that tested the effects of output, it was predicted that engaging in the SR task, in particular, would contribute to the improvement of their OEI test-performances in the posttests. Contrary to the original expectation, not only SR group but also the other two groups similarly improved their OEI test-performances in the posttests. Therefore, the results of the study failed to show beneficial effects of SR on learners' implicit knowledge development by comparing with the other instructional condition. A careful re-examination of the detailed research design and the procedures of the OEI test revealed an

important methodological recommendation and agenda for future research. Considering the potential practice effect of the OEI test, the results of the present study indicated the importance of including a test-only comparison group in the research design of effect-of-L2-instruciton studies. Furthermore, the unexpected results also shed light on some future research agenda regarding the use of the OEI test.

Another major finding of the study was the usefulness of explicit grammar explanation in form-focused instruction. Especially when the purpose of the instruction is to develop learners' explicit knowledge, providing explicit grammar explanation may be the most efficient and effective instructional strategies even if the target linguistic form is more complex, marked forms (e.g., RC [OPREP]). At the same time, it can also be said that the differential effects showed in the results of the OEI test and the FITB test support the criticism made by Doughty (2003, 2004) on use of measuring instruments used in the previous effect-of-instruction studies. As shown by the results of the study, the effects of different types of form-focused instruction greatly differ depending on whether the test can measure learners' implicit knowledge or just explicit knowledge. As Ellis (2015) claimed that the primary goal of L2 instruction must be developing learners' implicit knowledge, it is valuable to investigate more issues in ISLA using a valid and reliable measure that can draw learners' implicit knowledge.

Limitations of the Present Study and Directions for Future Research

It must be noted that this study was limited by several important methodological problems that need to be considered in future research. First of all, the major limitation was the small sample size. In this study, 33 Japanese EFL university students were assigned to one of the three instructional groups (SR, EGE, and Comparison), each of which consisted of eleven (or ten) students. Due to the small sample size for each group, it was very difficult to make

convincing arguments with quantitative results. Even though inferential statistics were used to supplement the findings from descriptive statistics, it is crucial to examine the topics and the research questions asked in this study by increasing the number of participants in future research.

Another major limitation of the study was found out after analyzing the results of the learners' test-performances. Contrary to the initial intentions, the comparison group also improved their test-performances significantly in the posttest sessions. However, it was impossible to identify whether their gains could be attributed to the input that they were exposed to in the treatment sessions or a possible test-practice effects caused by taking the OEI and FITB tests multiple times (or a combination of both). Therefore, in future research, especially in the type of research that uses the OEI test, including a test-only comparison group is crucial. In addition to this, it is also important to investigate whether and how much test-practice effect the OEI test potentially have in relation to some possible mediating variables, such as types of target linguistic form, the amount of learners' preexisting knowledge on the target linguistic form, and the number of test-items included in an OEI test.

Another issue also concerns with actual learning that occurred during the instructional treatments. Numerous studies explored the effects of explicit grammar explanation through learners' test-scores. It is possible that not all learners in the EGE group understood equally the explicit grammar explanations provided by their teacher. Even though the gains made by the EGE group may have been interpreted as their successful understanding of the explicit explanation provided during the instructional treatments, it also needs to be considered and checked the degree of learners' understanding of explicit explanation in future research.

It also needs to be noted here that the OEI and FITB tests used in this study were used as measures for assessing learners' implicit and explicit knowledge development. However, these

two types of measures are still not perfect measure of either type of knowledge, but rather are aimed to draw more on one type of knowledge than the other. For example, if a leaner possesses relatively high implicit knowledge on the target linguistic form, the learner can use the knowledge in the FITB test even though the test was used with the intention to draw L2 leaners' explicit knowledge. As Suzuki and Dekeyser (2015) argued, the learner may be able to use their highly automatized explicit knowledge in the OEI test even though the test was conducted with time constraints and focus on meaning facilitated by the use of the belief questionnaires. Therefore, the interpretation of the results needs to be made carefully.

Last but not the least, the time-gap between the immediate posttest and delayed posttest were also one the limitation of the study. Due to the researcher's scheduling conflict, the delayed posttest was conducted two weeks after the final instructional treatment. Considering the complex, time-consuming processes of the development of L2 implicit knowledge, two-week time-gap may be too short to claim that the results of the delayed posttest reflect long-term effects of the instructional treatments. Since developing implicit knowledge takes time, it is an important question to examine whether and how effects of form-focused instruction lead to the development of implicit knowledge over time. This question also needs to be further explored in future research with a longitudinal research design.

APPENDICES

APPENDIX A:
THE READING PASSAGES FOR THE INSTRUCTIONAL TREATMENTS

The first instructional treatment

Effective English Learning and Second Language Acquisition

Part 1

"I have studied English for eight years, but I can't speak English." This is a typical comment that many Japanese learners give. In the Japanese educational system, English is taught for eight years from junior high school to university. Despite many years of learning, why is it still difficult for many Japanese learners to use English? How can they learn English more effectively?

You can find some hints to the answers of these questions if you look at the findings of Second Language Acquisition (SLA) studies. SLA is a scientific field that studies how people learn a second language after acquiring their first language. Since the late 1960s, many SLA researchers have conducted research in which they studied the mechanisms of language learning. Based on the discoveries from SLA studies, an essential element in language acquisition is receiving 'a large amount of input.' In SLA, all language information that learners get from reading and listening is called input. Without input, it is impossible to acquire a language. For example, when you grow a plant, water and sunlight are the essential sources for growth. To drive a car, you have to put gas in your car to make it run. Similarly, the main source that helps second language development is input. All SLA researchers agree on the importance of input in language learning.

Looking back on your own English learning, how much input have you received inside and outside your English classrooms? Dr. Yasuhiro Shirai, a professor at Case Western Reserve University, comments in his book, "One of the biggest problems that many English teachers need to deal with is the lack of input in English classrooms." Another important question is what kind of input best helps your second language learning?

The second instructional treatment

Part 2

Input is the most important element in second language learning. Thus, it is essential to increase the amount of input in your learning through reading and listening. At the same time, the quality of input is also important. Based on the findings of SLA studies, what kind of input is necessary to learn a second language effectively?

SLA studies have shown that the input needs to be understandable. If you spend a lot of time reading or listening to things that you cannot understand, it is not helpful in learning a language. However, if the input is too easy, you can't move to the next level which you are aiming for. Therefore, input needs to be understandable but also a little bit challenging. In SLA, this type of input is called "comprehensible input." When Professor Shirai was a high-school English teacher in Japan, he taught his class based on SLA theories and gave a lot of

comprehensible input in his classes. After one year, his students' test scores dramatically improved. His students' deviation value (*hensachi*) increased 10 points in one year.

Another important factor which you need to care about for effective language learning is content familiarity. It is essential that the content of input be related to your own background and interests. If the content is interesting or familiar to you, you will read or listen to it more carefully and repeatedly. Also, the background knowledge that you have on the topic helps you to comprehend the input. If you like baseball, reading passages that are about baseball is a good way to get relevant input.

As discussed above, you can get many hints that help your learning from SLA research findings. As a compass shows you the right direction when you get lost in the woods, SLA theories guide you on your second language learning journey. Let's enjoy learning English and expand your infinite potential with SLA!

The third instructional treatment

Getting the most out of your study abroad

Study abroad is often viewed as a magical language learning opportunity. Some people naively expect that they can be a native-like English speaker after studying in an English-speaking country for one year or so. Is this scenario typically achieved by most English learners after completing their study abroad?

Despite these expectations, SLA literature shows that many students who experienced one-semester or one-year study abroad are often disappointed by the actual improvement that they made. Even if the person felt his/her improvement, objective proficiency tests, such as TOEFL or IELTS, do not often reflect as much improvement as the person felt. There are several pitfalls that prevent much improvement in study abroad.

First of all, students studying abroad do not usually receive as much input as they assume. In English-speaking countries, native speakers and university teachers tend to speak much faster than L2 learners can understand the message. As a result, the total amount of comprehensible input that they actually process is very limited. Another pitfall is related to the first one. Since L2 learners cannot fully understand other people's talk, their interactional mode tends to be entirely passive, such as just following other people's conversations and listening to university lectures. Also, learners do not often stretch out their current proficiency level because many situations in which learners need to use English can be dealt with some fixed formulaic expressions. For example, when ordering something at Starbucks, they just say, "Can I get a grande caramel Frappuccino?" "Here." "That's all, thanks." However, to develop L2 proficiency, learners need to construct sentences without relying too much on these fixed expressions so as to automatize the knowledge of grammar, vocabulary, pronunciation, etc.

To get the most out of a study abroad experience, a well-known SLA researcher named Robert Dekeyser points out the importance of systematic preparation for study aboard at home. If learners have not reached a certain level at which they can listen and speak without too much

struggle, they will miss the opportunity to receive a large amount of input and engage in output practice in real interactions. Based on implications from SLA research, learners can get some hints on where, when and how to engage in a good practice in their study abroad experience?

APPENDIX B:
DETAILED DESCRIPTION AND PRE-READING QUESTIONS ASKED IN THE FIRST PHASE

Session 1

First Paragraph

1. What's the problem described in the first paragraph?
2. How can you find some hints for effective English learning?

Second Paragraph

1. How can you find some hints for effective English learning?
2. What's the most important things in second language learning?
3. What's input?
4. What do SLA researchers agree on?

Third Paragraph

1. What did Professor Shirai say?
2. How much input have you received in your English classrooms?

Session 2

First Paragraph

1. What's the most important element in SLA?
2. How can you get input?
3. A large amount of input is a key to SLA. What's also important?

Second Paragraph

4. What's that?
5. What is "understandable but challenging input" called in SLA?

Third Paragraph

6. What's another important factor?
7. If you like baseball, what kind of reading and listening materials are helpful?

Session 3

First Paragraph

1. How do many language learners view study abroad?
2. What do they expect?

Second Paragraph

3. Is this true?
4. Why? What's the first pitfall?
5. What's another pitfall?
6. Entirely passive → Why?
7. To develop L2 proficiency, what's important?

Third Paragraph

8. What does the SLA researcher points out?

APPENDIX C:
THE VOCABULARY GLOSSES USED IN THE INSTRUCTIONAL TREATMENTS

First Session -- Vocabulary List

K-3 Words or Lower frequency words

- effective 効果的な
- effectively 効果的に
- hints ヒント
- acquire 習得する
- conduct ～を行う
- mechanism メカニズム、原理
- essential 必須の
- element 要素
- source 源、出どころ
- important 大切な
- professor 教授
- classroom クラスルーム

Second Session -- Vocabulary List

K-3 Words or Lower frequency words

- element 要素
- essential 不可欠な
- effective 効果的な
- aim 目指す
- comprehensible 理解可能な
- theory 理論
- factor 要因
- content 内容
- relevant 関連した
- hint ヒント
- expand ～を広げる
- potential 潜在能力

Third Session -- Vocabulary List

K-3 Words or Lower frequency words

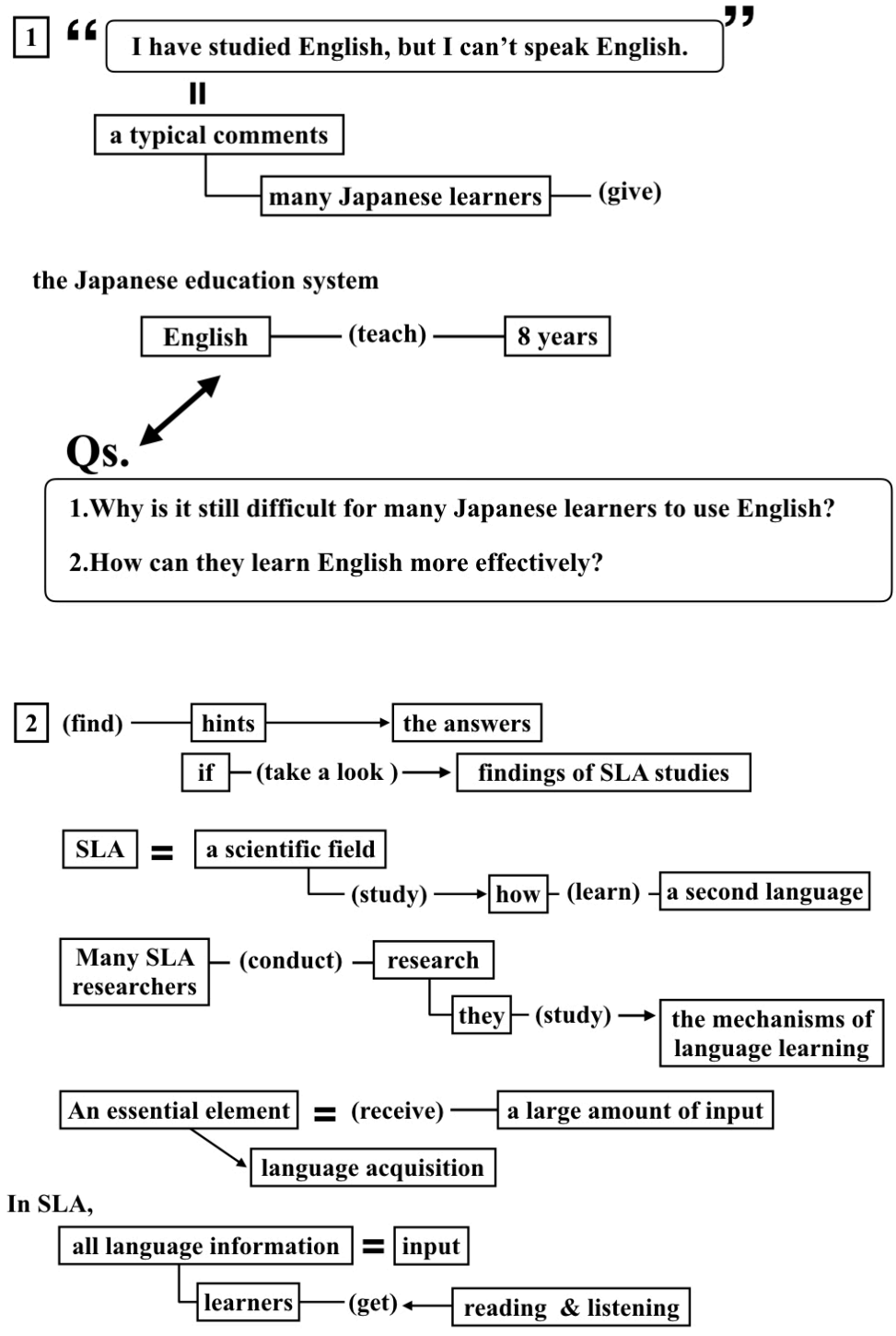
- naively 単純に
- achieve 達成する

- SLA literature　SLA 研究論文
- Semester　セメスター、学期
- objective　目標の
- proficiency　（言語）能力
- reflect　～を反映する
- pitfall　落とし穴
- interactional mode　インタラクションの形式
- passive　受け身
- lectures　講義、レクチャー
- formulaic expressions　熟語、セットの表現
- complex　複雑な
- construct　構成する、作る
- importance　大切さ
- implication　示唆

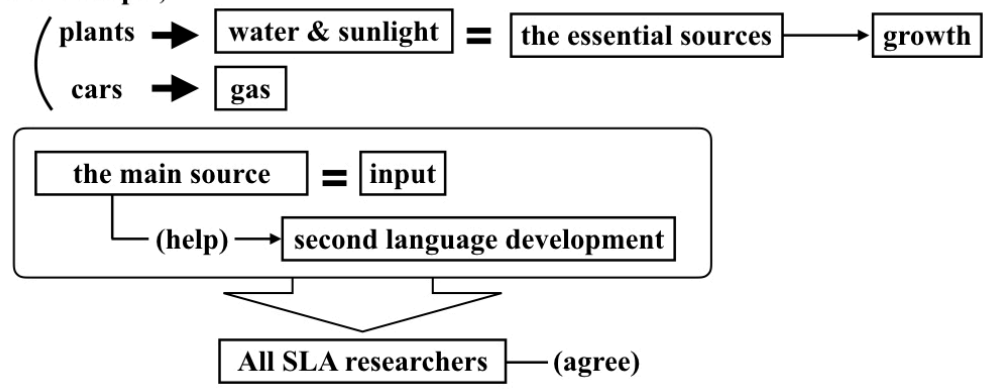
APPENDIX D:
CONCEPT MAPS FOR THE SR TASK

Session 1

Effective English Learning and Second Language Acquisition



For example,



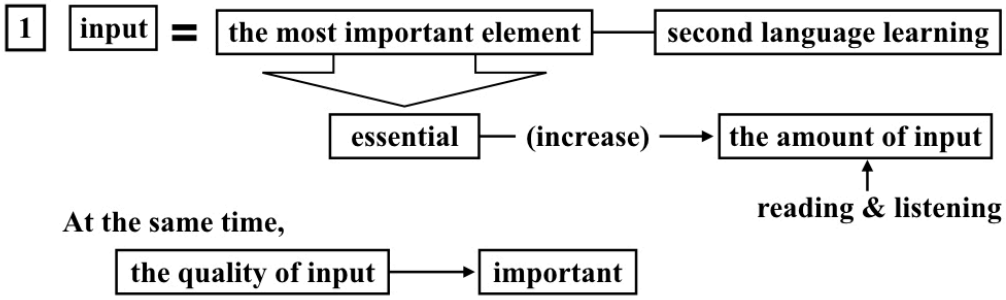
3 Looking back → English learning

how much input ← (receive) — your English classroom

Dr. Yasuhiro Shirai → "One of the biggest problems = the lack of input
English teachers — (deal with)"

What kind of input best helps your second language learning?

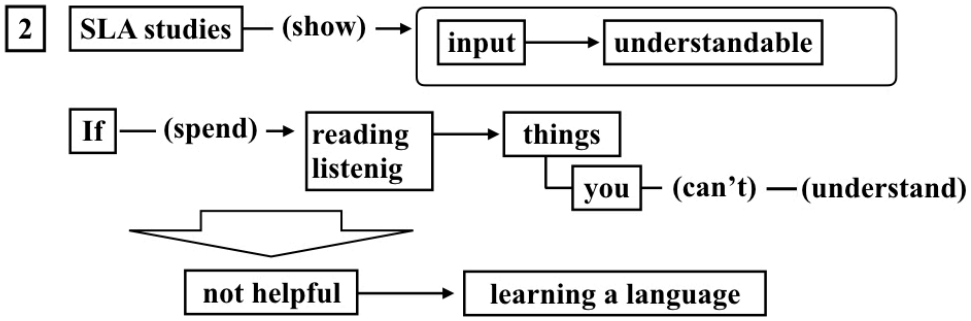
Session 2

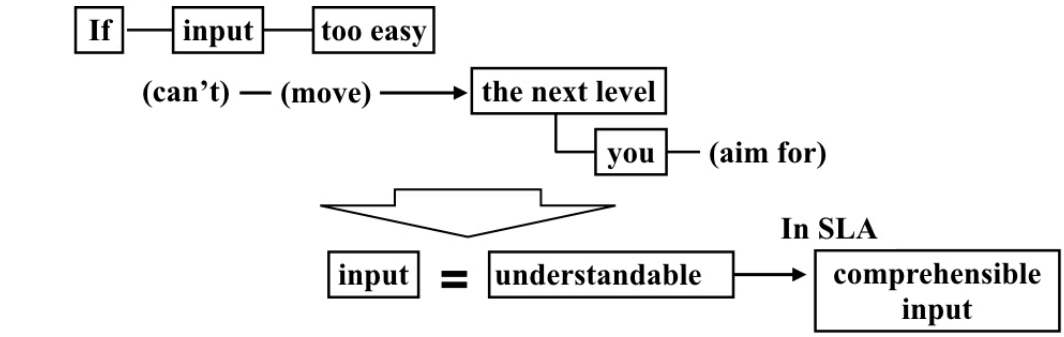


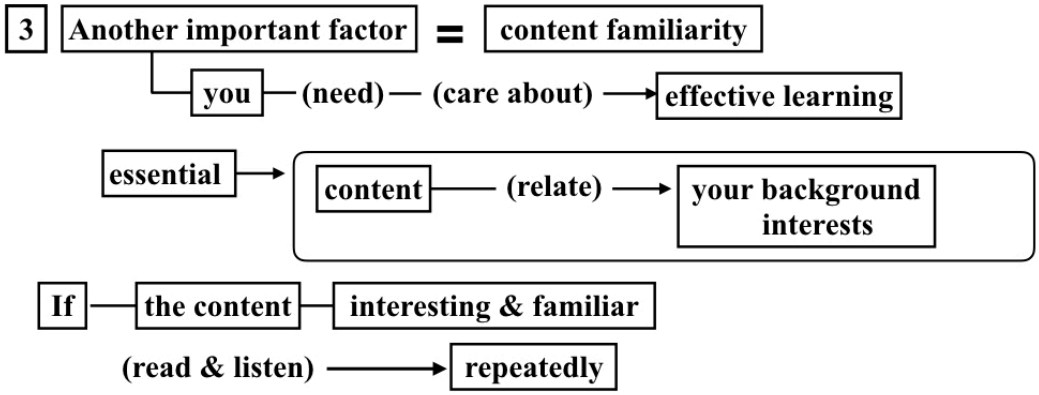
At the same time,

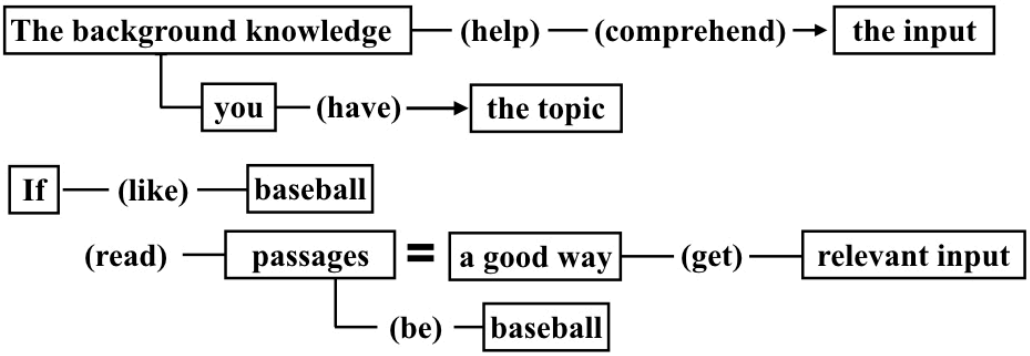
the quality of input → important

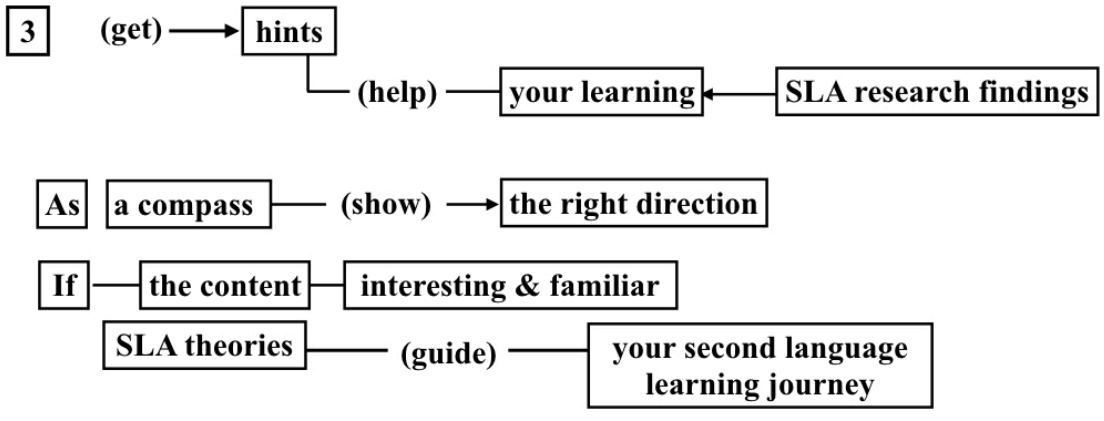
Q. What kind of input is necessary to learn a second language effectively?




If
input
too easy
(can't)
(move)
the next level
you
(aim for)
In SLA
input
=
understandable
comprehensible input

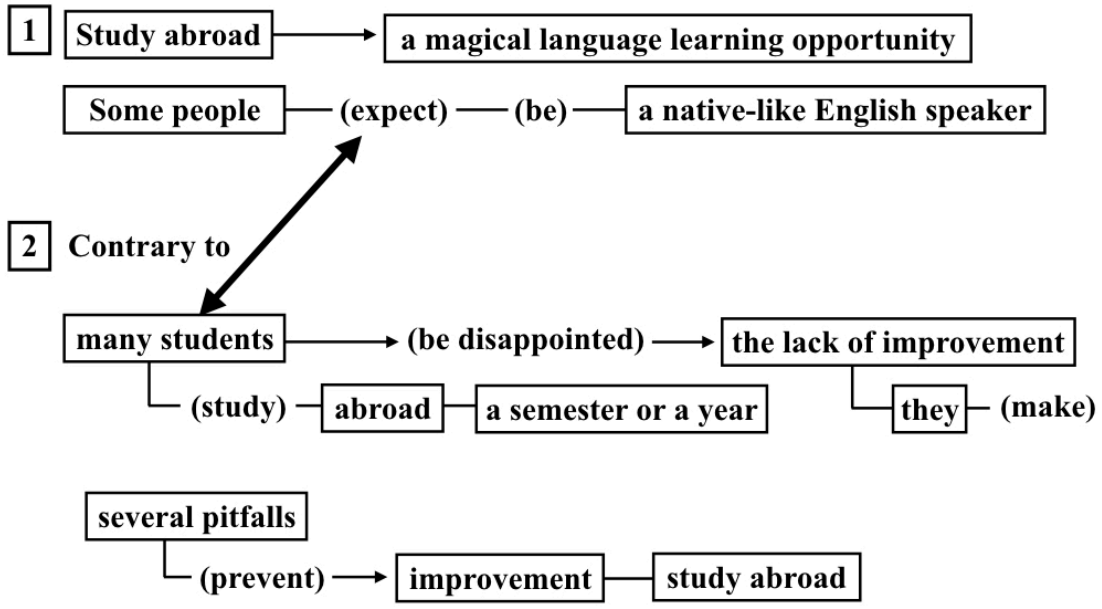

3
Another important factor
=
content familiarity
you
(need)
(care about)
effective learning
essential
content
(relate)
your background interests
If
the content
interesting & familiar
(read & listen)
repeatedly


The background knowledge
(help)
(comprehend)
the input
you
(have)
the topic
If
(like)
baseball
(read)
passages
=
a good way
(get)
relevant input
(be)
baseball


3
(get)
hints
(help)
your learning
SLA research findings
As
a compass
(show)
the right direction
If
the content
interesting & familiar
SLA theories
(guide)
your second language learning journey

Session 3

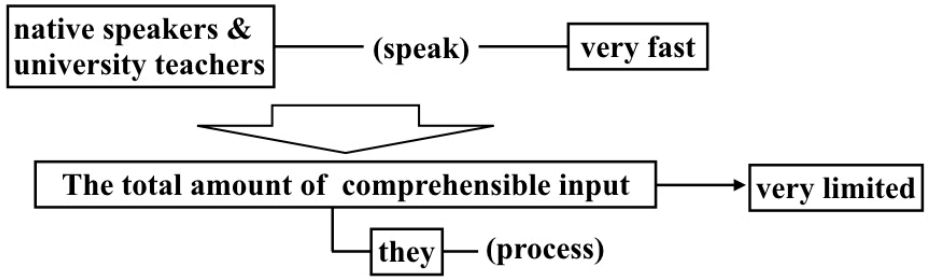
Getting the Most out of Your Study Abroad



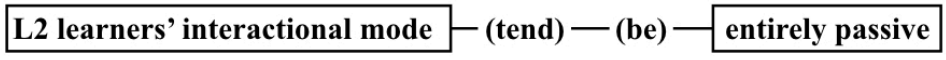
3

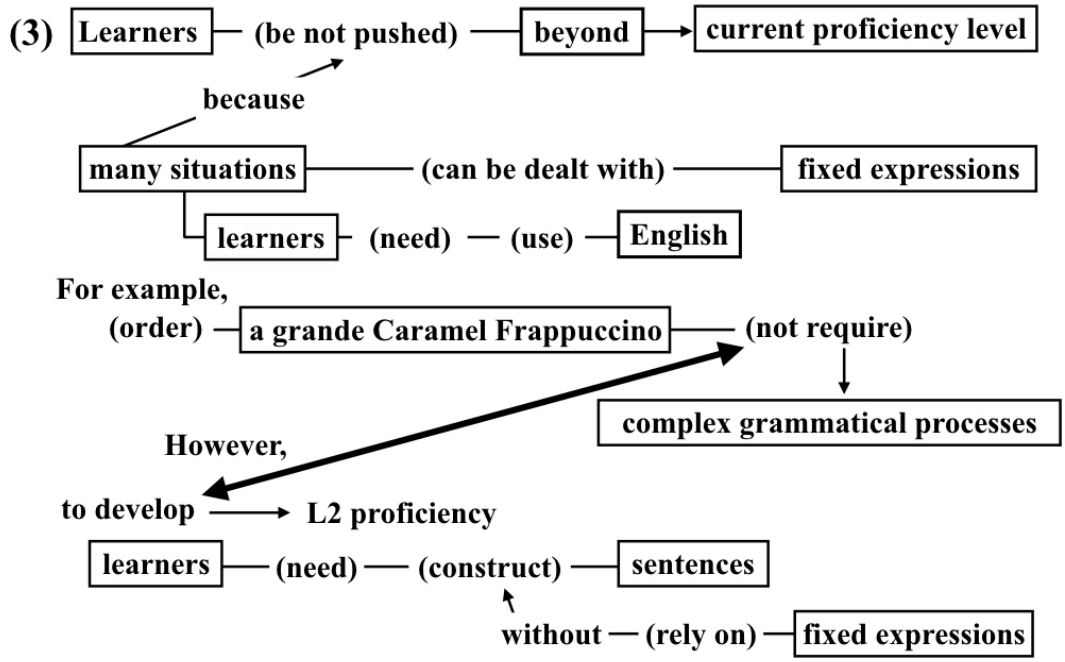
(1) First of all,

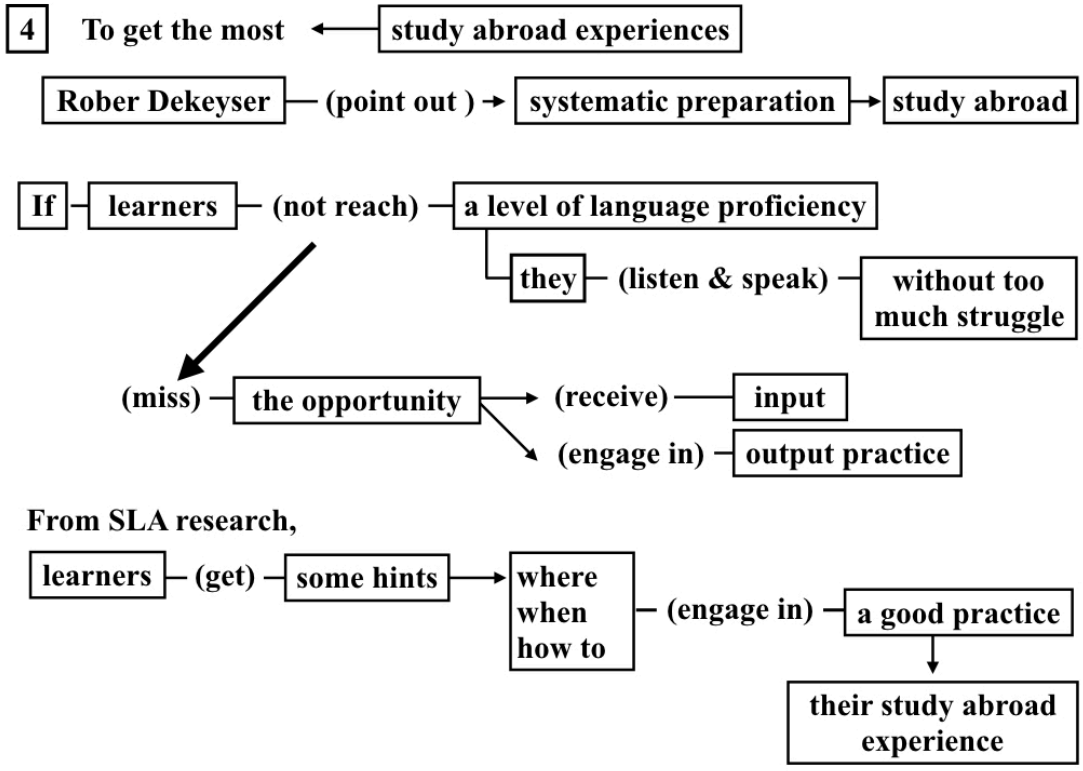
In English speaking country,



(2) Another pitfall




(3)
Learners
(be not pushed)
beyond
current proficiency level
because
many situations
(can be dealt with)
fixed expressions
learners
(need)
(use)
English
For example,
(order)
a grande Caramel Frappuccino
(not require)
complex grammatical processes
However,
to develop
L2 proficiency
learners
(need)
(construct)
sentences
without
(rely on)
fixed expressions


4
To get the most
study abroad experiences
Rober Dekeyser
(point out)
systematic preparation
study abroad
If
learners
(not reach)
a level of language proficiency
they
(listen & speak)
without too much struggle
(miss)
the opportunity
(receive)
input
(engage in)
output practice
From SLA research,
learners
(get)
some hints
where
when
how to
(engage in)
a good practice
their study abroad experience

APPENDIX E:
THE SUMMARY OF THE EXPLICIT GRAMMAR EXPLANATION

First Session

文法説明：３つの関係代名詞 – 主格、直接目的語、前置詞の目的語の関係代名詞

関係代名詞 who, which, that を使うと、名詞（先行詞）を後ろから節で説明（修飾）することができます。つまり、関係代名詞は、文の中のある名詞（人やもの）について、それがどういう人なのか、またはどういうものなのかを、さらに詳しく説明することによって意味を限定していく働きをしています。

❖ 説明される名詞（先行詞）が人の場合は、who または that が用いられ、先行詞が人以外の場合は、which または that が用いられます。

1. 主格の関係代名詞

→ 関係代名詞が、後ろで説明される節の中で主語の役割をします。

Similarly, **the main source** ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ is input.
主な源 ← < **that** helps second language development >
S V O

SLA is **a scientific field**
科学的な領域 ← < **that** studies how people learn a second language
S V O after acquiring their first language >.

2. 直接目的語の関係代名詞

→ 関係代名詞が、後ろで説明される節の中で直接目的語の役割をします。

In SLA, **all language information** ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿ is called input.
全ての言語情報 ← < **(that)** learners get -- from reading and listening >
S V (O)

“I have studied English for eight years, but I can’t speak English.”
This is **a typical comment**
典型的なコメント ← < **(that)** many Japanese learners give -- > .
S V (O)

3. 前置詞の目的語の関係代名詞

→ 関係代名詞が、後ろで前置詞の目的語の役割をします。

"**One of the biggest problems** ———————————— is the lack of input in English classrooms."
一つの大きな問題 ↰ < **(that)** many English teachers need to deal **with --** >
S V (O)

Since the late 1960s,
many SLA researchers have conducted **research**
研究 ↰ < **in which** they studied the mechanisms
S V O
of language learning -- >.
(in research)
(O)

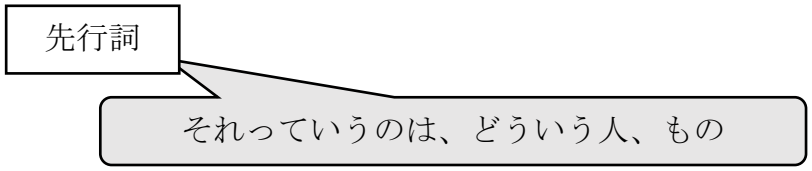
❖ 前置詞の目的語になっている場合、その前置詞も一緒に関係代名詞節の先頭に置くことができます。この場合、who, that を使うことはできず、which（人の場合は whom）を使い、関係代名詞を省略することもできません。ただ、前置詞が関係代名詞節の後ろに残った場合は、関係代名詞に that を用いることができます。

Second Session

文法説明：３つの関係代名詞 – 主格、直接目的語、前置詞の目的語の関係代名詞

関係代名詞 who, which, that を使うと、名詞（先行詞）を後ろから節で説明（修飾）することができます。つまり、関係代名詞は、文の中のある名詞（人やもの）について、それがどういう人なのか、またはどういうものなのかを、さらに詳しく説明することによって意味を限定していく働きをしています。

❖ 説明される名詞（先行詞）が人の場合は、who または that が用いられ、先行詞が人以外の場合は、which または that が用いられます。



4. 主格の関係代名詞

→ 関係代名詞が、後ろで説明される節の中で主語の役割をします。

If you like baseball, **reading passages** ———————— is a good way to get relevant input.
読むための文章 ↰ < **that are about baseball** >
S V

You can get **many hints** ———————— from SLA research findings.
多くのヒント ↰ < **that help your learning** >
S V O

5. 直接目的語の関係代名詞

→ 関係代名詞が、後ろで説明される節の中で直接目的語の役割をします。

In SLA, **all language information** ———————————————— is called input.
全ての言語情報 ↩ < **(that)** learners get -- from reading and listening >
S V (O)

The background knowledge ———————— helps you to comprehend the input.
↩
背景知識 < (**that**) **you have** -- **on the topic** >
S V (O)

6. 前置詞の目的語の関係代名詞

→ 関係代名詞が、後ろで前置詞の目的語の役割をします。

Another important factor ———————————————— is content familiarity.
もう一つの重要な要因 ↩ < **which** you need to care **about** -- for effective language learning
S V (O)

If the input is too easy, you can't move to **the next level**
次のレベル ↩ < **which** you are aiming **for** -- .
S V (O)

❖ 前置詞の目的語になっている場合、その前置詞も一緒に関係代名詞節の先頭に置くことができます。この場合、who, that を使うことはできず、which（人の場合は whom）を使い、関係代名詞を省略することもできません。ただ、前置詞が関係代名詞節の後ろに残った場合は、関係代名詞に that を用いることができます。

Third Session

文法説明：３つの関係代名詞 – 主格、直接目的語、前置詞の目的語の関係代名詞

関係代名詞 who, which, that を使うと、名詞（先行詞）を後ろから節で説明（修飾）することができます。つまり、関係代名詞は、文の中のある名詞（人やもの）について、それがどういう人なのか、またはどういうものなのかを、さらに詳しく説明することによって意味を限定していく働きをしています。

❖ 説明される名詞（先行詞）が人の場合は、who または that が用いられ、先行詞が人以外の場合は、which または that が用いられます。

7. 主格の関係代名詞

→ 関係代名詞が、後ろで説明される節の中で主語の役割をします。

Many students ———————————— are often disappointed by the lack of improvement.
多くの生徒 ↰ < **who** study abroad for a semester or a year >
S V

There are **several pitfalls**
いくつかの落とし穴↰ < **that** prevent improvement in study abroad. >
S V O

8. 直接目的語の関係代名詞

→ 関係代名詞が、後ろで説明される節の中で直接目的語の役割をします。

The total amount of **comprehensible input** ———————————— is very limited.
全ての言語情報 ↰ < **that** they actually process -- >
S V (O)

Many students are often disappointed by the lack of **improvement**
成果 ↰ < **that they make** -- >.
S V (O)

9. 前置詞の目的語の関係代名詞

→ 関係代名詞が、後ろで前置詞の目的語の役割をします。

Many situations ———————— can be dealt with by using fixed formulaic expressions.
多くの状況 ↰ < **in which** learners need to use English -- >
S V O (O)

If learners have not reached **a level of language proficiency**
言語能力のレベル↰ < **at which** they can listen and speak without
S V
too much struggle -- **>,**
(O)
they will **miss the opportunity to** receive a large amount of input and to engage in output practice in real interactions.

❖ 前置詞の目的語になっている場合、その前置詞も一緒に関係代名詞節の先頭に置くことができます。この場合、who, that を使うことはできず、which（人の場合は whom）を使い、関係代名詞を省略することもできません。ただ、前置詞が関係代名詞節の後ろに残った場合は、関係代名詞に that を用いることができます。

APPENDIX F:
REVISED EXAMPLE SENTENCES FOR THE TWO TESTS

Oral Elicited Imitation (OEI) Test

Novel Statements (24 Statements)

SU: 6 items (3 subject & 3 object)

The matrix subject position (x 3)

1. The baseball league that has the strongest players is Major League.
2. Large cars that use a lot of gas are not popular in Japan.
3. The only animal that can use language is a human being.
4. A prefecture that doesn't have any mountains is Yamanashi.

The matrix object position (x 3)

5. Ichiro is a baseball player who belongs to an American team.
6. President Obama is the first US president who visited Hiroshima.
7. In Japan, there are many convenience stores that sell vegetables.
8. Kyoto is a city that has many beautiful places.

DO: 6 items (3 subject & 3 object)

The matrix subject position (x 3)

9. The cars that American people want is Toyota.
10. One thing that you can do to stay healthy is not to eat McDonalds.
11. A school subject that most elementary school students like is math.
12. Knowledge that you learn in school helps you in your future.

The matrix object position (x 3)

13. Nature has many things that human beings cannot understand.
14. Donald Trump is a person that most American people respect.
15. Spanish is the language that people in Brazil speak.
16. The best food that you should eat in Osaka is takoyaki.

OPREP: 6 items (3 subject & 3 object)

The matrix subject position (x 3)

17. The house that the US president lives in is the Black House.
18. A learning goal that most Japanese students aim for is to enter Tokyo University.
19. The best application with which you can share your pictures is Instagram.
20. The best teachers that children can learn from are their grandparents.

The matrix object position (x 3)

21. Tokyo University is the school that the Prime Minister Abe studied at.
22. Ueno zoo is the first zoo that a panda arrived at.
23. Global warming is an issue that all Japanese people must think about.
24. Overwork is a problem that Japanese society should think about.

The Untimed Fill-in-the-Blank Test
Novel Sentences (24 Sentences)

SU: 6 items (3 subject & 3 object)

The matrix subject position (x 3)

1. (The best application that navigates your car to your destination) is Google Map.
2. (The best sightseeing place that has a lot of beautiful beaches) is Okinawa.
3. (The school subject that helps you most in your life) is English.
4. (The country that has the largest population in the world) is China.

The matrix object position (x 3)

5. A local train is (the type of train that stops at every station) along the way.
6. Albert Einstein is (a scientist who made a lot of great inventions).
7. Toyota is (a company that has several factories) in the US.
8. Starbucks is (the cafe that sells the best coffee) in the world.

DO: 6 items (3 subject & 3 object)

The matrix subject position (x 3)

9. (One skill that all Japanese high school students should learn) is computer skills.
10. (The best transportation that you can use to go) to Hokkaido is plane.
11. (The biggest mistake that human beings made) in history is World War II.
12. (The natural disaster that most Japanese people fear) is the earthquake.

The matrix object position (x 3)

13. Hawaii is (the place that most Japanese people want to visit).
14. Oda Nobunaga is (a person that most Japanese people respect).
15. Recent NHK broadcasts (many TV programs that young people can enjoy).
16. *Furoshiki* is (a cloth that Japanese people use to wrap things).

OPREP: 6 items (3 subject & 3 object)

The matrix subject position (x 3)

17. (The best strategy with which you can get higher test-scores) is to not sleep.
18. (A problem that Japanese government have to deal with) is aging society.
19. (One important goal that Japanese university students aim for) is to study abroad.
20. (The first thing that school students should pay attention to) is to avoid plagiarism.

The matrix object position (x 3)

21. Children's grades are (things that most parents care about).
22. If you want to be rich, becoming a doctor is (something that you should aim for).
23. Hitotsubashi University is (the university that Mr. Mikitani graduated from many years ago).
24. School teachers have (many things that they have to think about).

REFERENCES

REFERENCES

Adams, R. (2003). L2 output, reformulation and noticing: Implications for interlanguage development. *Language Teaching Research, 7,* 3, 347-376.

Ammar, A., & Lightbown, P. M. (2005). Teaching marked linguistic structures- More about the acquisition of relative clauses by Arab learners of English. In A. Hausen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 167-198). Berlin: Mouton de Gruyter.

Bardovi-Harling, K. (1995). The interaction of pedagogy and natural sequences in the acquisition of tense and aspect. In F. Eckman, D. Highland, P. Lee, J. Mileham, & R. Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 151-168). Mahawa, NJ: Lawrence Erlbaum Associates.

Brown, H. D. (2000). *Principles of language learning and teaching*. White Plains, NY: Longman.

de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning,* 46, 3, 529-555.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), The handbook of second language acquisition (pp. 312–348). Oxford, UK: Blackwell.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language learning*, *55*(S1), 1-25.

DeKeyser, R. (2016). Of moving targets and chameleons-Why the concept of difficulty is so hard to pin down. *Studies in Second Language Acquisition, 38(2),* 353-363.

Dekeyser, R. (2017). Knowledge and skill in ISLA. In S, Loewen, & M, Sato (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition* (pp. 15-32). New York: Taylor & Francis.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*, 431–469.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Eds.), *Cognition and second language instruction*, (pp. 206-257). New York: Cambridge University Press.

Doughty, C. (2003). Instructed SLA: Constraints, Compensation, and Enhancement. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition*, (pp.256-310), Oxford: Blackwell.

Doughty, C. (2004). Effects of Instruction on Learning a Second Language: A Critique of Instructed SLA Research. In B. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form meaning connections in second language acquisition* (pp. 181-202). Mahwah, US:

Taylor and Francis.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in second language acquisition*, *27*(2), 141-172.

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied linguistics*, *27*(3), 431-463.

Ellis, R., Loewen, S. D., Elder, C., Erlam, R., Philp, J., & Reinders, H. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. (Volume 42). Mulilingual Matters.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, *27*, 464–491.

Erlam, R. (2013). Explicit knowledge and grammar explanation in second language instruction. In C. Chapelle (Eds.), *The encyclopedia of applied linguistics.* Oxford, England: Wiley Blackwell. doi:10.1002/9781405198431.wbeal.0404

Erlam, R., and Akakura, M. (2016). New developments in the use of elicited imitation. In A. Mackey, & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research into second languages* (pp. 94-111). Taylor and Francis.

Erlam, R., Loewen, S., & Philp, J. (2009). The roles of output-based and input-based instruction in the acquisition of L2 implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching*, 241-261.

Fotos, S. (1998). Shifting the Focus from Forms to Form in the EFL Classroom. *ELT Journal*, *52*(4), 301-7.

Gass, S. (1988). Integrating research areas: A framework for second language studies1. *Applied linguistics*, *9,* 2, 198-217.

Gass, S. (1997). *Input and interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hamilton, R. (1994). Is implicational generalization unidirectional and maximal? Evidence from relativization instruction in a second language. Language Learning, 44, 123-157.

Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, *11*(4), 459-479.

Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing, 21(4)*, 332-347.

Izumi, S. (2000). *Promoting noticing and SLA: An empirical study of the effects of output and*

input enhancement on ESL relativization. Unpublished Ph.D. dissertation, Georgetown University.

Izumi, S. (2002a). Output, input enhancement, and the Noticing Hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition 24*: 541-77.

Izumi, S. (2002b). EFL education in japan from the perspective from the second language acquisition research, *Sophia Linguistica, 49,* 3-14.

Izumi, S. (2003a). Comprehension and production processes in second language learning. *Applied Linguistics,* 24, 2, 168-196.

Izumi, S. (2003b). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning* 53, 285-323.

Izumi, S. (2013a). Noticing and L2 Development: Theoretical, Empirical, and Pedagogical Issues. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka, (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 37-50). Honolulu: University of Hawai'i, National Foreign Language Resource Center.

Izumi, S. (2013b). Focus on form (FonF). In P. Robinson (Eds.), *The Routledge encyclopedia of second language acquisition*. Routledge.

Izumi, S. (2016). *Focus on form to CLIL no eigo zyugyo*. [*Focus on form and CLIL in English classroom*]. Tokyo: ALC Press.

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition?. *Tesol Quarterly*, *34*(2), 239-278.

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition, 21*, 421–452.

Izumi, S., and Izumi, Y. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *The Canadian Modern Language Review, 60(5),* 587-609.

Keenan, E., and Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*, 63-99.

Kuno, S. (1974). The position of relative clauses and conjunctions. *Linguistic Inquiry, 5*, 117-136.

Larson–Hall, J. (2016). A guide to doing statistics in second language research using SPSS and R. New York: Routledge.

Li, W., and He. X., (2017, October). *The effectiveness of written output on promoting L2 learners' attention: An eye-tracking study*. Poster session presented at Second Language

Research Forum 2017, Columbus, OH.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam/Philadelphia: John Benjamins.

Long, M. H. (1998). Focus on Form in Task-based Language Teaching. *University of Hawai'i Working Papers in English as a Second Language, 16 (2),* 35-49.

Long, M., and Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.) *Focus on form in classroom second language acquisition* (pp. 15-41). New York: Cambridge University Press.

Loewen, S. (2013). Instructed second language acquisition. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2716-2718). Malden, MA: Blackwell Publishing.

Loewen, S. (2015). *Introduction to instructed second language acquisition.* New York: Rutledge.

Loewen, S., & Sato, M. (2017). *The Routledge Handbook of Instructed Second Language Acquisition*. Taylor & Francis.

Morgan-Short, K., & Bowden, H. W. (2006). Processing instruction and meaningful output-based instruction: effects on second language development. *Studies in Second Language Acquisition*, *28*(1), 31-65.

Muranoi, H. (2006). *Dainigengosyutokuriron kara mita koukatekina eigo gakusyuho shidoho [Effective English learning and teaching methods from the perspective of second language acquisition].* Tokyo: Taisyukan.

Muranoi, H. (2007a). Output practice in the L2 classroom. In R. Dekeyzer (Eds.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 51-84). New York: Cambridge University Press.

Muranoi, H. (2007b). Focus on form through guided summarizing and EFL learners' interlanguage development. *Journal of the English institute, 33,* 15-59. Tohoku Gakuin University.

Muranoi, H. (2009). *Inputo to autoputo wo tsunagu kyokasyo chushin no zyugyo: kyoshitsu SLA kenkyu karano shisa* [MEXT-aouthorized textbook-based classroom, connecting the processes from input to output: Implications from classroom SLA]. The English Teachers' Magazine (Eigokyoiku), 57(12), 31-33.

Pinemann, M. (1998). *Languae processing and second language development: Processability theory*. Amsterdam: Benjamins.

Russell, V. (2014). A closer look at the output hypothesis- The effect of pushed output on noticing and inductive learning of the Spanish future tense. Foreign Language Annals, 47(1), 25-47.

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, *63*(3), 595-626.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Spada, N. and Tomita, Y. (2010). Interaction between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60,* 2, 263-308.

Spada, N., Shiu, J. L. J., and Tomita, Y. (2015). Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning*, *65*(3), 723-751.

Suga, K. (2016). *The effects of story-retelling in PCPP instruction: Output practice in meaning oriented focus on form.* Unpublished M.A. thesis, Sophia University, Tokyo.

Suzuki, Y., and DeKeyser, R. M. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*, 860–895.

Swain, M. (1985). Communicative competence: Some role of comprehensible input and comprehensible outputting tis development, In S, Gass and C. Medden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, Mass.: Newbury House.

Swain, M. (1995). There functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). Cambridge University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-481). New York: Cambridge University Press.

Swain, M. and Lapkin, S. (1995). Problems in output and cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371-91.

Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition. *Language learning*, *56*(2), 319-385.

Uggen, M. S. (2012). Reinvestigating the noticing function of output. *Language learning*, *62*(2), 506-540.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7,* 133-161.

Williams, J., and Evans, J. (1998). What kind of focus and on which forms?. In C. Doughty & J.

Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139-155). Cambridge University Press.

Williams, J., and Paciorek, A. (2016). Indirect tests of implicit linguistic knowledge. In A, Mackey, & E, Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research into second languages* (pp. 25-42). Taylor and Francis.