

THE EFFECT OF PRE-TASKS ON THE QUALITY OF TIMED-WRITING

A Dissertation Proposal
Submitted
to the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY

By
Naoko Hosoda
December 2023

Examining Committee Members

David Beglar, Advisory Chair, Teaching and Learning
Tomoko Nemoto, Teaching and Learning
Edward Schaefer, External Member, Ochanomizu University
Tamara Swenson, External Member, Osaka Jogakuin University

©
Copyright
2023

by

Naoko Hosoda
All Rights Reserved

ABSTRACT

This investigation consisted of two studies related to a timed-writing task. The first study was an investigation of the effect of three pre-tasks on the quality of timed-writing focusing on five variables: syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility. The second study concerned the longitudinal effect of engaging in timed-writing on the development of syntactic complexity, writing fluency, and lexical diversity.

The participants were 84 university students studying English at a university in Japan in which the timed-writing task was implemented as a part of class activity for over nine months. The study adopted a Latin squares design and participants in the experimental group completed 12-minute timed-writing tasks with three pre-task conditions: outlining, oral rehearsal, and reading. For the outlining condition, the participants completed 5 minutes of individual planning before completing the timed-writing task. For the oral rehearsal pre-task, they rehearsed the plan they completed during the planning session orally in pairs before completing the timed-writing task. For the reading pre-task, the participants read an article related to the timed-writing topic before the planning and timed-reading sessions. Of 18 timed-writing sessions, the participants had chance to engage in each of the three pre-tasks six times, and the texts written at Time 1, Time 3, Time 4, and Time 6 were used as data. The participants in the comparison group only took the pretest and posttests. The pretest was conducted and before the intervention session, Posttest 1 was conducted after three months, and Posttest 2 was conducted nine months after the intervention period.

The first phase of Study 1 was focused on the analysis of the quantitative data. To compare the effect of three pre-tasks on five variables—syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility—the texts were converted to quantitative data using text analysis software that made the data ready for statistical analysis. The texts were scored by raters for content and comprehensibility and used for analysis.

The second phase of Study 1 was focused on the analysis of the qualitative data. Based on the findings from the quantitative data, stimulated recalls were conducted with three students. The purpose of these sessions was to examine what the participants did during the pre-task and planning sessions and how their thinking affected the quality of their final timed-writing products. The post-course questionnaire and additional text analyses were also conducted to supplement the findings of the quantitative data.

Study 2 concerned the longitudinal effect of implementing the timed-writing tasks on the development of syntactic complexity, writing fluency, and lexical diversity. The pretest and two posttests the participants in the experimental group and comparison group completed were compared.

The results revealed that none of the three pre-tasks impacted the syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility measures of the immediate writing task to a statistically significant degree. One possible reason is the participants' English proficiency level. As suggested in a previous study by Johnson et al. (2012), the participants' English proficiency must reach a certain level in order to benefit from a pre-task that reduces the working memory burden. The participants in this study

were CEFR A2 level, which is categorized as a basic language user, so their English proficiency was perhaps not developed enough to fully benefit from the pre-tasks.

The results of the questionnaire indicated that nearly 50% of the participants found reading pre-task most helpful because it was effective in terms of idea generation, and learning vocabulary and grammar. Their comments were supported by the descriptive statistics of content and lexical diversity scores, which outperformed outlining group. About 40% of the participants found oral rehearsal was most helpful because it was effective for idea generation and organizing the idea. However, the comments included both positive and negative ones. How much they can benefit from oral rehearsal pre-task is more influenced by individual differences compared with other two pre-tasks. Although some participants found outlining task useful, only 10% of the participants found it was most effective.

Regarding the longitudinal effect of implementing the timed-writing activity, the current study showed that it had a significant effect on developing writing fluency. After continuing timed-writing activities for nine months, greater gains were observed for the experimental group than for the comparison group. For syntactic complexity and lexical diversity, there only significant difference between the pretest and two posttests for both groups was mean length of clause. However, because the growth was significant for the experimental group and the comparison group, it was not merely the effect of timed-writing. There was no time effect or group differences on any of three lexical diversity measures. The results of the study support the developmental order of CFL measures suggested in previous studies that propose that they develop in the order of writing fluency, syntactic complexity and lexical diversity.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to all those who have supported and guided me throughout this journey of pursuing my Ph.D. degree. Without their advice, help, and encouragement, I would not have been able to finish this dissertation.

First and foremost, I am profoundly grateful to my advisor, Dr. David Beglar, for his continuous guidance and mentorship. His expertise, patience, and dedication have been invaluable in completing this work.

Second, I would like to express my gratitude to Dr. Tomoko Nemoto. She devoted a great deal of time to help me with the statistical analyses and I could not have completed this work without her support. I am also thankful to the members of my dissertation committee, Dr. Edward Schaefer and Dr. Tamara Swenson. Their insightful feedback has greatly enriched the quality of this research.

I extend my appreciation to my colleagues and friends who participated in this research. My special gratitude goes to Junko Yamaai, Yukie Saito, Kumiko Katase, and Junko Baierschmidt, who devoted their time during their busy schedules to this study. Their advice and encouragement helped me to stay motivated to complete this study.

I would like to express my gratitude to the members of my doctoral cohort and my family. Their continuous encouragement and support have been a constant source of strength. Last but not least, I would like to acknowledge all the students who participated in this study.

Naoko Hosoda

October 2023

TABLE OF CONTENTS

	PAGE
ABSTRACT.....	iii
ACKNOWLEDGMENTS	vi
LIST OF TABLES	xii
LIST OF FIGURES	xiv
CHAPTER	
1. INTRODUCTION.....	1
Key Terminology	1
The Background of the Issue	1
Statement of the Problem.....	3
Purposes of the Study.....	4
The Audience for the Study	6
Delimitations.....	7
Organization of the Study	8
2. REVIEW OF THE LITERATURE.....	9
Kellogg’s Writing Model and Working Memory	9
Intervention Studies of Pre-Tasks and Writing.....	15
Pre-Task Planning	15
The Effect of Individual and Collaborative Planning on Writing	23
Rehearsal.....	31
Reading for Writing	36
Factors That Affect Writing Performance	41
Meaning-Focused Writing and Writing Development	43

Gaps in the Literature.....	49
Purposes of the Study.....	50
Research Questions.....	52
3. METHODS.....	53
Context of the Study	53
Participants.....	55
Research Design.....	56
Instrumentation	58
Timed-Writing Task.....	58
The Writing Tests.....	60
Pre-Task Conditions	61
Outlining.....	61
Oral Rehearsal Pre-Task.....	61
Reading Pre-Task	62
Post-Course Questionnaire.....	64
Procedures.....	65
Pretest	66
Timed-Writing Tasks	68
Posttests.....	70
Interview Sessions.....	71
Analyses.....	72
Complexity, Fluency, and Lexical Diversity Measures.....	72
Syntactic Complexity.....	73
Writing Fluency	74
Lexical Diversity.....	75

Holistic Measurements	75
Content	76
Comprehensibility	76
Interviews and Post-Course Questionnaire	79
Additional Text Analysis on Lexical Items.....	80
4. PRELIMINARY ANALYSIS	81
The Overview of the FACETS Analysis	81
Comprehension Quizzes for Reading Passages	84
5. RESULTS.....	86
The Effect of the Three Pre-Tasks on Timed-Writing	86
The Roles of Pre-Tasks	100
Post-Course Questionnaire	100
Interviews	102
Outlining.....	107
Oral Rehearsal	109
Reading.....	112
Additional Text Analysis of Lexical Items	119
The Effect of Timed-Writing on Writing Development.....	119
Syntactic Complexity and Lexical Diversity	120
Mean Length of T-Unit.....	120
Dependent Clause per Clause	122
Mean Length of Clause.....	124
Writing Fluency: Standard Word Units	125
Lexical Diversity.....	127
Lexical Density	127

Moving-Average Type Token Ratio	129
Measure of Textual, Lexical Diversity	131
6. DISCUSSION	133
The Effect of Three Pre-Tasks on Syntactic Complexity, Writing Fluency, Lexical Diversity, Content, and Comprehensibility	133
The Roles of Pre-Tasks	141
Outlining	141
Oral Rehearsal.....	142
Reading	144
The Longitudinal Effect of Timed-Writing on Writing Development	145
Research-Oriented Findings.....	150
Pedagogical Implications	152
7. CONCLUSION	157
Summary of the Findings.....	157
Limitations of the Study.....	160
Suggestions for Future Research	161
Concluding Comments.....	163
REFERENCES	164
APPENDICES	
A. OUTLINING SHEET	175
B. COMPLETED WORKSHEET FOR OUTLINING	176
C. PASSAGES FOR THE READING GROUP	177
D. POST-COURSE QUESTIONNAIRE	191
E. INFORMED CONSENT FORM: JAPANESE VERSION	192
F. INFORMED CONSENT FORM: ENGLISH TRANSLATION.....	194
G. TIMED-WRITING PROGRESS CHART	197

H. INTERVIEW QUESTIONS	198
I. TIMED-WRITING SAMPLE TEXTS	199

LIST OF TABLES

Table	Page
1. Descriptive Statistics for the TOEIC Scores of the Participants by Class in Study ...	56
2. Descriptive Statistics for the TOEIC Scores by Group	56
3. Timed-Writing Topics	59
4. Lexical Composition, Flesch-Kincaid Readability Scores, and Standard Word Units of the Reading Passages.....	63
5. Timing of the Pre-Task Treatments for Each Task Condition	64
6. Schedule of the Study for the Experimental and Comparison Groups.....	67
7. Measures Used for Complexity, Fluency, and Lexical Diversity Analyses.....	73
8. Six-Point Holistic Scale.....	77
9. Rasch Statistics for the Six-Point Evaluation Criteria.....	83
10. Category Statistics for the Six-Point Rating Scale	84
11. The Percentage of Correct Answers for Each Comprehension Quiz.....	85
12. Descriptive Statistics for the Syntactic Complexity Measures for Each Pre-Task....	87
13. Descriptive Statistics for Standard Word Units for Each Pre-Task	93
14. Descriptive Statistics for the Lexical Diversity Measures for Each Pre-Task.....	96
15. Descriptive Statistics for Content and Comprehensibility for Each Pre-Task.....	99
16. Most Helpful Timed-Writing Pre-Tasks and the Reasons	101
17. The Percentage of Word Family Similarity Rate Between the Reading Passage and Timed-Writing for Each Pre-Task Group.....	120
18. Descriptive Statistics for Mean Length of T-Unit for the Experimental and Comparison Groups.....	121
19. Descriptive Statistics for Dependent Clause per Clause for the Experimental and Comparison Groups.....	123
20. Descriptive Statistics for Mean Length of Clause for the Experimental and Comparison Groups.....	125

21. Descriptive Statistics for Standard Word Unit for the Experimental and Comparison Groups.....	127
22. Descriptive Statistics for Lexical Density for the Experimental and Comparison Groups	128
23. Descriptive Statistics for Moving-Average Type Token Ratio for the Experimental and Comparison Groups	130
24. Descriptive Statistics for Measure of Textual, Lexical Diversity for the Experimental and Comparison Groups	131

LIST OF FIGURES

Figure	Page
1. Kellogg's Writing Model.....	11
2. The Design of the Study	58
3. Wright Map for Content and Comprehensibility	82
4. Changes in Mean Length of T-Unit.....	88
5. Changes in Dependent Clauses per Clause.....	88
6. Changes in Mean Length of Clause.....	89
7. Changes in the Number of Standard Word Units per 12 Minutes	92
8. Changes in Lexical Density	97
9. Changes in the Moving-Average Type Token Ratio (MATTR)	97
10. Changes in the Measure of Textual Lexical Diversity (MTLD).....	98
11. Changes in the Mean Length of T-units for the Experimental and Comparison Groups	122
12. Changes in the Dependent Clause per Clause for the Experimental and Comparison Groups	123
13. Changes in the Mean Length of Clause for the Experimental and Comparison Groups	125
14. Changes in the Standard Word Units for the Experimental and Comparison Groups	127
15. Changes in the Lexical Density for the Experimental and Comparison Groups	129
16. Changes in the Moving-Average Type Token Ratio for the Experimental and Comparison Groups	130
17. Changes in the Measure of Textual, Lexical Diversity for the Experimental and Comparison Groups	132

CHAPTER 1

INTRODUCTION

Key Terminology

In this study, the terms below are defined as follows.

- Syntactic complexity: Grammatical sophistication expressed in the length or the number of a particular syntactic unit
- Writing fluency: The number of words or units a writer can write within a given time period
- Lexical diversity: The variety of the words used in writing. Lexical richness.
- Content: The quantity and quality of ideas or concepts in the participants' written texts (Kuiken & Vedder, 2017).
- Comprehensibility: The degree of effort that is required to understand writing (Kuiken & Vedder, 2017).

The Background of the Issue

Due to globalization and the 2020 Tokyo Olympics, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT), announced a new English Education Reform Plan in 2013. The plan suggested that upper secondary school English teachers should “nurture the ability to understand abstract contents for a wide variety of topics and the ability to fluently communicate with English speaking persons” and include activities that require “high-level linguistic ability” (MEXT, 2013, p. 1) such as presentations, debates, and negotiations. This plan shows that unlike traditional approaches to English education in Japan that have emphasized teaching

reading, vocabulary, and grammatical structures, students are expected to acquire the ability to successfully express their ideas in English in written and oral forms.

Following this plan, a major revision in the standardized college entrance examination created by the National Center for University Entrance Examinations is planned. For the foreign language section, the Center is considering the use of English proficiency examinations such as the Test of English for Academic Purposes (TEAP), Test of English as a Foreign Language (TOEFL), and International English Language Testing System (IELTS) to measure students' writing and speaking proficiency. This reform in the college entrance examination will bring about changes in the English curriculum in Japanese secondary schools. In order to prepare students for university entrance examinations, teachers are expected to spend more time to foster students' writing skills, specifically, the ability to write paragraphs or essays in testing situations.

English education in Japan has traditionally been focused on grammar and reading instruction rather than the productive skills of speaking and writing, and of the four major language skills—listening, reading, speaking, and writing—writing receives the least attention in secondary school English courses (Mochizuki et al., 2015).

Evidence of this situation was provided by Ohi (2016), who administered a questionnaire to 129 Japanese junior and senior high school English teachers and reported that of the four major language skills, writing received the least attention in secondary school English courses. When writing is taught, Japanese teachers often focus on linguistic features such as vocabulary and grammar and only 22% of the teachers who responded reported that they teach paragraph or essay writing.

Traditionally, English writing education in Japan has emphasized grammatical accuracy

and as a result, insufficient attention has been paid to writing beyond the sentence level. Although syntactic accuracy is an important aspect of writing, students also need to be able to generate ideas, access appropriate linguistic forms to express those ideas, and produce written output under time pressure.

In addition to the lack of attention to writing instruction, Ohi (2016) reported that many Japanese teachers are not confident about teaching writing; thus, they feel that they need to learn more about how to do so. When asked to indicate the main problem in teaching English writing, 43% of the teachers stated that they needed a better understanding of English writing and how to teach it, and 35% responded that it is necessary to develop activities that promote the acquisition of English writing skills. Therefore, more empirical studies of the impact of classroom interventions on Japanese students' writing skills should be conducted.

Statement of the Problem

This study was focused on three problems. The first problem is that few researchers have investigated the effects of writing pre-tasks and their effect on writing development. To date, the majority of studies focused on writing pre-tasks have concerned pre-task planning. These studies include planning time (Rostamian et al., 2018; Seyyedi et al., 2013; Tabari, 2017) with and without time pressure (Ellis & Yuan, 2004) and solitary vs. collaborative planning (Dobao, 2012; Doe & Figueroa, 2015; Tavakoli & Rezazadeh, 2014). Compared with oral studies, the variety of studies that have been focused on the effect of pre-tasks on L2 writing is small; thus, it is necessary to conduct additional studies to clarify how various types of pre-tasks affect writing performances, namely syntactic complexity, writing fluency, lexical diversity, and

content and comprehensibility. For example, although opinion sharing or reading a passage are common pre-writing activities in foreign language classrooms, few researchers have focused on how these activities affect writing fluency and writing quality. In addition, the previous literature does not include studies with low-proficiency participants (Polio, 2017). By targeting participants with relatively low English proficiency, I hope to add new insights to the understanding of the effects of pre-writing tasks on writing.

The second problem is that even though some researchers have provided evidence that pre-writing tasks exert a positive effect on writing compared with a non-planning condition, it is not clear how planning activities have helped participants improve their performances. As Ellis and Yuan (2004) indicated, more studies are needed to explore the relationship between process and product by investigating what L2 writers do under different pre-task planning conditions.

The third problem is the lack of longitudinal investigations of how fluency-first, meaning-focused writing activities contribute to writing development. Some studies exist in this field, but of those studies, few were conducted for more than one semester (Nguyen, 2015; Nitta & Baba, 2014) and the inclusion of a comparison group has been rare. In addition, mixed results have been reported; thus, more studies are needed to clarify the situation.

Purposes of the Study

The first purpose of this study was to examine the effect of three types of pre-writing tasks—outlining, oral rehearsal, and reading—on syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility. Numerous studies on the

effect of planning on oral performance using the CAFL (complexity, accuracy, fluency, lexical diversity) framework have shown that different types of strategic planning, such as solitary planning, teacher-led planning, and interactive planning in pairs or groups, affect oral performances differently (Ellis, 2009; Foster & Skehan, 1999; Kawauchi, 2005; Mochizuki & Ortega, 2008). Although there are similarities between oral production and writing production, Ellis and Yuan (2005) stated that they differ in terms of time pressure and learners' ability to monitor their output. While speakers can experience time pressure when producing speech, writers can read, think, and edit their output as they write because they are usually under relatively little time pressure. Therefore, the first purpose is to investigate how three types of pre-writing activities—outlining, oral rehearsal, and reading—affect timed-writing performances.

The second purpose is to examine how the participants make use of each type of pre-writing task and how the tasks influence the final written products. Previous researchers have shown that planning positively affects linguistic variables such as syntactic complexity and fluency, but it has not been clear what cognitive processes contributed to those changes (Ellis & Yuan, 2004; Seyyedi et al., 2013). I therefore examined this issue by conducting stimulated recall interviews in which the participants recalled what they were thinking while engaging in the pre-writing tasks. By conducting a mixed method study, which has not been attempted by the majority of researchers in this area, it is possible to better understand the relationship between what learners do during pre-task planning and their final written products.

The third purpose was to explore the longitudinal effect of implementing timed-writing activities on syntactic complexity, writing fluency, and lexical diversity. By

comparing the results between three intervention groups and a comparison group as well as with those of the previous studies, this study adds new insights to the understanding of the fluency-first, meaning-focused writing activities.

The Audience for the Study

This study can benefit two groups in the field of second language acquisition (SLA) and foreign language education. The first group is researchers conducting similar writing studies in the field of SLA and related fields. As mentioned above, few researchers have investigated the effect of pre-task planning on syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility in writing with a mixed method design. This study can provide researchers with new insights into the efficacy of three writing pre-tasks and the cognitive processes L2 learners use while engaging in pre-writing tasks by providing information concerning what factors contributed to the quantitative results.

The second audience is writing teachers, material designers, and curriculum developers who design writing courses. A large number of researchers have investigated the effect of planning (e.g., with or without planning; pre-task planning vs. online planning; with or without time pressure) on writing, but pre-writing tasks other than planning have not received much attention. In particular, no researchers have investigated the influence of pre-reading activities on written production. As Ohi (2016) showed, teachers hope to acquire more information on how to teach writing courses and the effect of various pre-writing activities. Findings on the effect of different types of pre-writing tasks on the final written products will help instructors and curriculum developers select tasks to use in their lessons.

Furthermore, the findings will enable material designers to develop better materials by giving them more options for pre-writing tasks that positively affect students' writing. This study provides data concerning the effects of three pre-writing tasks: outlining, oral rehearsal, and reading.

Delimitations

The first delimitation is related to two internal factors, the participants' L1 and their English language proficiency levels. The participants' first language is Japanese, a language that differs considerably from English in numerous ways, including orthography, lexis, syntax, morphology, and discourse structure. Therefore, the results of this study might differ from previous studies of native speakers of English as well as studies whose participants' L1 is genetically related to English. Another internal factor concerns the students' English proficiency level. The participants in this study had low English proficiency (TOEIC 370–400); thus, the results are not generalizable to second language learners with higher English proficiency. This study was conducted with EFL learners who struggle with transforming their ideas into written English. In contrast, learners with sufficient writing fluency will not benefit from pre-writing activities as they can complete writing tasks without support from the pre-tasks used in this study.

The third delimitation concerns the topics and genre of the writing used in this study. To minimize the effect of topic familiarity, general topics related to university students' daily lives such as traveling, their home, and shopping were selected. The genre of the writing was an opinion paragraph. Thus, the results of this study might not be transferable to more complex topics that require specialized knowledge or to different genres such as narrative or storytelling.

Organization of the Study

Chapter 2, Review of the Literature, is made up of three major sections: Kellogg's (1996) writing model and working memory, intervention studies on planning and writing, and tasks that can facilitate writing performance. At the end of the chapter, the gaps in the literature, the purposes of the study, and the research questions are presented. In Chapter 3, Methods, I describe the context of the study, the participants, the instrumentation, the research design, the procedures, and the analytical tools used in the study. In Chapter 4, Preliminary Analyses, I present the results of the FACETS analyses. In Chapter 5, Results, I first present the results for the quantitative analyses, and then the qualitative analyses. In Chapter 6, Discussion, I combine the findings of both the quantitative and qualitative results and discuss their implications and relationship to previous studies. I also discuss the pedagogical implications of the results. In Chapter 7, Conclusion, I summarize the findings, consider the limitations of the study, make suggestions for future research, and provide concluding comments.

CHAPTER 2

REVIEW OF THE LITERATURE

In this study, I investigated the effectiveness of three types of pre-writing tasks—outlining, speaking, and reading—on writing quality, which was operationalized as complexity, fluency, lexical complexity, content, and comprehensibility. In the following sections, I first review Kellogg’s (1996) writing model, which was created based on Levelt’s (1989) speech model, and then explain why writing is a challenging skill to develop for second language learners using Skehan’s (1998) Limited Attentional Capacity Model. I also consider how providing pre-writing tasks can help writers reduce the information processing burden. I then focus on three types of pre-writing tasks relevant to this study: strategic planning, rehearsal, and reading a passage. After that, I review intervention studies in which researchers have investigated how pre-writing tasks affect the complexity, accuracy, lexical diversity, and fluency of the participants’ writing compositions, and then introduce longitudinal investigations of how fluency-first, meaning-focused writing activities contribute to writing development. I then highlight the gaps in the literature addressed by this study, present the purposes of the study, and list the research questions that guided this study.

Kellogg’s Writing Model and Working Memory

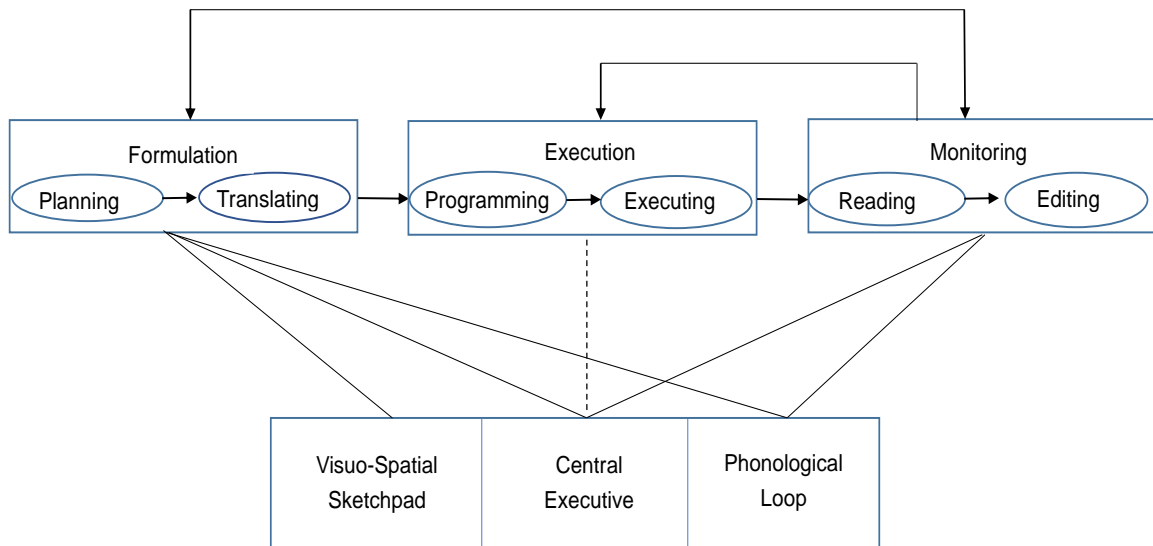
Writing is a complex activity that involves interactions between a wide range of cognitive processes. In the field of L1 writing studies, the importance of long-term memory and working memory has been emphasized by numerous researchers (Bereiter & Scardamalia, 1987; Flower & Hayes, 1980; 1981). Based on Levelt’s (1989)

influential speech model, Kellogg (1996) proposed a writing model. Although his model was originally developed to describe the process for L1 writing, it was selected as the basis for the current study because it explains the writing process in relation to working memory, and it has been referred to in many of the papers reviewed in this Chapter.

Figure 1 shows Kellogg's (1996) writing model and the resources of working memory used in the three stages of formulation, execution, and monitoring. The formulation stage involves setting goals, generating ideas, and translating them into linguistic forms. The execution stage comprises programming the motor system and executing muscle movement. Finally, in the monitoring stage, writers read what has already been written and edit organizational problems and linguistic errors. Each of the three stages consists of two substages, and the arrows indicate the flow of the information. For example, the output of the formulation stage is sent to the execution stage, and the output of the execution stage is sent to the monitoring stage. However, Kellogg (1996) stated that writing should not be interpreted as a linear process. Rather, the model shows that all three stages interact with each other and the process is recursive in nature. For instance, writers would notice that editing is necessary before the execution stage as well as the need to go back to the formulation stage to make necessary corrections. Kellogg also stated that the formulation of new text and monitoring the existing text can occur concurrently. The double-headed arrows and the backward arrow in the figure indicate this property. Kellogg (1996) proposed that all of

the stages are activated simultaneously as long as the demands placed on the central executive do not exceed its capacity.

Figure 1. *Kellogg's Writing Model*



From *The Science of Writing* (p.59), by C. M. Levy and S. Ransdell (Eds.), 1996, Taylor and Francis Group (<https://doi.org/10.4324/9780203811122>). Copyright 1996 by Taylor and Francis Group. Reproduced with permission of The Licensor through PLSclear.

Kellogg (1996) stated that each of the three stages of formulation, execution, and monitoring consists of two substages. The formulation stage consists of planning and translating substages. In the planning substage, writers set writing goals, generate ideas, and select the necessary information for the writing task. At this stage, output might be a propositional representation that can be translated into linguistic forms, but it can also appear as non-linguistic abstract images or feelings. In the next substage, translating, output from planning is translated into linguistic forms. In order to turn ideas into complete sentences, the selection of the appropriate lexical units, the

formation of a syntactic frame, phonologically representing the lexical units in the frame, and the transformation of phonemes into orthographic patterns must take place.

The second substage of Kellogg's (1996) writing model is the execution stage, which consists of two substages: programming and executing. In the programming substage, the output from the translation stage activates the program that controls appropriate parts of the motor system for executing the writing task. In the case of typing, the muscles required for moving the arm, hand, and fingers are activated while handwriting involves the selection of the appropriate letter sizes and allographic forms as well. In the executing substage, necessary motor units are recruited for muscle movement.

Finally, in the monitoring stage, writers read and edit the text they are producing. While reading, the subprocesses that control word recognition, sentence comprehension, and coherence are activated. In the editing stage, the writer's intentions and the written output are compared. If there is a mismatch, the text is edited or revised. Monitoring occurs not only after the executing substage but also during the formulation stage, and involves actions that range from noticing local grammatical errors to assessing paragraph organization.

Kellogg (1996) elaborated on the model by incorporating the three components that make up Baddeley's (1986) working memory model: the central executive, the phonological loop, and the visuospatial sketchpad. How these components relate to the three stages mentioned in the previous section is visualized in Figure 1. The central executive is responsible for the retrieval of information in long-term memory. It

controls attention and coordinates two subsidiary systems, the phonological loop and visuospatial sketchpad. Both the formulation and monitoring stages are connected to the phonological loop, whose role is to store and process auditory and verbal messages. When sentences are produced, writers silently recite them to themselves; thus, the phonological loop is activated together with the central executive while translating. The phonological representations of the words are stored in the phonological loop briefly to form more sophisticated syntactic units or before the information is sent to the execution stage. The phonological loop is also activated during the reading stage, one of the monitoring stages.

The visuospatial sketchpad, on the other hand, stores and processes visual and spatial information and is connected to the formulation stage in Figure 1. It is activated during planning with the central executive because writers plan by visualizing ideas, organizing, considering orthography, and thinking about the layout. The phonological loop and the visuospatial sketchpad can function alone, but their information processing capacity is limited. When the task demand is too high and they cannot handle processing independently, the central executive is recruited to help complete the task. The demand of the central executive on the execution stage is minimal unless writers are young children who have not had enough writing practice. That issue is represented with a dotted line.

All the processes and functions described above—formulation, execution, and monitoring—and the three working memory functions—the central executive, phonological loop, and visuospatial sketchpad—work simultaneously. For this reason,

writing is a highly cognitively demanding skill that places stress on the limited capacity central executive. According to Kellogg (1996), the cognitive demand of the formulation stage is the greatest because this system places a burden on all three working memory resources (i.e., the central executive, phonological loop, and visuospatial sketchpad). The cognitive demand of the execution stage is the least demanding.

He also stated that planning, translating, and reviewing compete for limited working memory resources. Kellogg (2001) stated that when the demand of planning is low, spare resources are distributed equally to planning, translating, and reviewing. On the other hand, when the motor execution demand is low, the resources are distributed to reviewing, followed by translating. However, the effect on planning was not significant. Kellogg indicated that the result supported the hypothesis that planning, translating, and reviewing share common working memory resources, but how they are distributed depends on the particular case in question.

Kellogg (1996) noted that pre-task planning can enhance writing performance. Online planning and translating compete for working memory resources; thus, by separating the planning and translating stages, more working memory resources are available for the translating stage while composing text, which leads to higher fluency and potentially more complex grammar. The effect of pre-task planning on syntactic complexity is not clearly stated in his writing model.

Similar ideas where speaking is concerned were proposed by Skehan (1998) in his Limited Attentional Capacity model. Skehan proposed that L2 learners have limited

working memory capacity, and trade-off effects occur when they focus their attentional resources on one aspect of the target language. For example, if they focus on syntactic complexity, then syntactic accuracy often suffers because as sentences become longer, there are more opportunities to make errors. To reduce the information processing burden, the effective application of pre-task planning promotes better performance by L2 learners (Foster & Skehan, 1996; Ortega, 1999; Yuan & Ellis, 2003). For second language learners who have been exposed to relatively small amounts of the target language, the formulation and execution stages present difficulties because these skills have not been proceduralized or automatized. As a result, producing the second language requires considerable cognitive effort for many L2 writers.

Intervention Studies on Pre-Tasks and Writing

Pre-Task Planning

As mentioned above, writing is a cognitively demanding skill especially for lower-proficiency L2 writers. Thus, the implementation of pre-writing tasks that facilitate language processing and idea generation can potentially improve writing performances. One of the most well-known pre-writing tasks for promoting higher-quality writing is planning. Kellogg (1990) stated that writing a brief outline during pre-writing eases the burden of composing, and as a result, positively affects both writing fluency and writing quality.

Bygate (2001) defined planning as “mental rehearsal of the message and its formulation prior to communication” (p. 28). In line with this idea, numerous studies on the effects of planning on oral performances have been conducted using the CALF

framework. Though small in number compared to speaking studies, some studies (Ellis & Yuan, 2004; Rostamian, et al., 2018) have been focused on writing. Even though the idea that pre-task planning has a positive effect on CALF measures is supported by these empirical studies, there are mixed results considering the effect of planning time on each CALF measure.

The effect of three planning conditions on writing quality was explored by Ellis and Yuan (2004). The participants were 42 university students studying English in China who were assigned to one of three conditions: (a) no-planning (- planning, + time pressure), (b) pre-task-planning (+ planning, + time pressure), and (c) on-line-planning (- planning, - time pressure). The number of syllables produced per minute and the number of dysfluencies operationalized as words that were crossed out or changed were used to measure writing fluency. The results showed that the pre-task planning groups produced more syllables than the no planning and on-line planning groups ($F = 9.80, p < .01$) and produced fewer dysfluencies ($F = 3.74, p < .03$), suggesting that pre-task planning is useful for improving writing fluency. Pre-task planning also resulted in greater syntactic variety compared to a no planning group ($F = 9.05, p < .001$). There were no statistically significant differences across the three groups for syntactic complexity ($F = 2.71, p = .079$), lexical variety ($F = 0.18, p < .84$), or syntactic accuracy (Error-free clauses: $F = 3.39, p = .059$; correct verbs: $F = 1.98, p = .15$), but the results indicated that there was an advantage for the online planning group regarding syntactic accuracy. The authors reported that although they could not conclude that pre-task planning is more effective than online planning due to the difference in the total

time spent on the task, pre-task planning enhanced writing fluency and syntactic complexity because it reduced the burden of the central executive and facilitated translating. It also gave the learners greater confidence during the task performance. The researchers concluded that pre-task planning could improve writing fluency and syntactic complexity while online planning might positively affect syntactic accuracy and suggested the need for both types of planning in order to produce the best performance.

Based on Ellis and Yuan (2004), Rostamian et al. (2018) investigated the effect of planning condition on syntactic complexity, syntactic accuracy, and writing fluency, and how their participants' limited attentional resources were allocated at each stage of the writing process: planner/proposer, translator, and evaluator/reviser. Sixty intermediate-proficiency Iranian EFL learners were divided into four groups: (a) no planning (NP: - planning, + time pressure), (b) on-line planning (OLP: - planning, - time pressure), (c) pre-task planning (PTP: + planning, + time pressure), and (d) pre-task and on-line planning (PTOLP: + planning, - time pressure). Their findings were similar to that of Ellis and Yuan (2004). The pre-task planning group had the highest syllable per minute score, $F = 4.54$, $p = .006$. and the two groups who completed pre-writing tasks—the pre-task planning and pre-task and on-line planning groups—scored higher than the on-line planning group, $F = 4.09$, $p = .011$, for clause length, indicating that pre-task planning enhanced syntactic complexity and writing fluency. On the other hand, the online planning group outperformed the pre-task groups on the syntactic accuracy measures of error-free clauses, $F = 6.55$, $p < .001$, and correct verb forms, $F =$

4.46, $p = .007$. Stimulated recall interviews indicated that comments about the cognitive processes involved in planning, translating, and evaluating were reduced at the time of writing for the pre-planning groups, while more comments about self-repairs were observed for the online planning group. The authors stated that these results explained the difficulty of drawing attention to all aspects of the language simultaneously and supported Skehan's (1998) Limited Attentional Capacity Model.

The impact of planning time on writing quality was examined by Tabari (2016), who examined the effect of pre-task planning and online planning on descriptive essays written by 78 EFL learners who were randomly divided into three groups: pre-task planning, online planning, and no planning. The participants in the pre-task planning condition performed better than those in the other groups in terms of writing fluency, $F = 16.28, p < .001$, while the participants in the no-planning group produced texts with greater lexical diversity. The online planning group outscored the pre-task planning group in terms of syntactic accuracy and lexical diversity, but there was no statistically significant difference. Tabari stated that the no-planning group outperformed the pre-task planning group on lexical diversity due to the trade-off effect between writing fluency and using a wider variety of lexis. In the pre-task planning condition, the learners focused on increasing production speed and the quantity of writing at the expense of lexical diversity. On the other hand, the no planning condition allowed the participants to use a greater variety of lexis. In addition, Tabari indicated that lexis should be evaluated separately from syntactic complexity, and four linguistic

dimensions—complexity, accuracy, lexis, and fluency—should be taken into account when analyzing L2 learners' writing performances.

The common findings from the studies reviewed above are that pre-task planning effectively promotes writing fluency, but it does not promote greater syntactic accuracy or lexical diversity. However, Ong and Zhang (2010) and Johnson et al. (2012) reported contrasting results. Ong and Zhang explored the effect of task complexity on writing fluency and lexical complexity of argumentative essays written by 108 EFL learners. There were three factors for task complexity: (a) planning time (extended pre-task, pre-task, free-writing, and control), (b) the availability of ideas and macro-structure (topic, ideas, and macro-structure given; topic and ideas given; and topic given), and (c) draft availability (draft available, no draft available). In the draft available condition, the participants were allowed to see their first drafts while those in the no draft available condition had nothing to refer to when completing their essays. The researchers found that the effect of planning time on one measure of writing fluency—the mean number of words produced per minute/total time spent on the task—significantly affected writing fluency, $F = 56.34, p < .001$, but the post hoc test showed that the fluency of the free-writing group (no pre-task planning) was greater than that of the pre-task ($p < .001$) and extended pre-task groups ($p < .001$). Two reasons were stated to explain the results. The first reason concerned the time spent writing. The on-line planning group had 30 minutes to plan and write, the extended pre-task group had 20 minutes to plan and 10 minutes to write, and the pre-task group had 10 minutes to plan and 20 minutes to write. The second reason is that the participants in the two pre-

task groups could have been engaging in online planning during the writing task. If that was the case, the authors stated, the pre-task planning conditions would have been more cognitively demanding than the free-writing conditions. The researchers suggested that pre-task planning can negatively impact fluency if planning time is too long or it is prolonged to the execution stage.

Similar to Ong and Zhang (2010), Johnson et al. (2012) did not find strong evidence to support that pre-task planning enhances writing fluency. They investigated the impact of five pre-task planning groups on writing. The participants, 968 advanced-proficiency Peruvian learners of English, were randomly assigned to five pre-task planning groups: (a) control (vocabulary exercise), (b) idea generation (brainstorming activity), (c) organization (writing an outline), (d) goal setting (answering questions such as “What is the purpose of writing this essay?,” “Who is the audience?,” and (e) organization + goal setting. There were no statistical differences in the total number of words produced by five pre-task groups, a finding that differed from previous studies on L1 writing fluency as well as with the results of Ellis and Yuan (2004), who reported a large impact on writing fluency. Johnson et al. (2012) indicated that one possible explanation is that writing is a recursive activity and the participants were engaged in online planning or editing during the timed-writing task. Thus, the benefit of pre-task planning was diluted by online planning and editing, which could explain the small impact on writing fluency.

Johnson et al. (2012) also stated that educational background and genre knowledge also contributed to the differing results. In Ellis and Yuan’s (2004) study, a

story telling task, which was not cognitively demanding and a genre the participants were familiar with, was used for the writing task. In addition, the authors of previous studies on outlining targeted university students writing in their first language. They had expertise in academic writing and genre knowledge, which could be another explanation for the contrasting results. Finally, the authors noted the importance of language proficiency. In order for writers to benefit from pre-task planning that reduces the working memory burden during the translating process, their language proficiency must reach a certain threshold. For low-proficiency students who struggle to produce the target language, pre-task planning might not ease their cognitive load during writing.

The findings on syntactic complexity are more complicated. Tabari (2017) and Rostamian et al. (2018) reported positive effects for pre-task planning on syntactic complexity. Tabari investigated the impact of planning time on the syntactic and lexical complexity of argumentative essays with 90 EFL learners divided into three groups: (a) pre-task planning, (b) online planning, and (c) no planning. He found that pre-task planning positively affected syntactic complexity ($F = 11.74, p < .001$, pre-task x no planning, $p < .001$, online x no planning $p < .002$) as well as syntactic variety ($F = 19.58, p < .001$; pre-task x no planning, $p < .001$, online x no planning $p = .130$, pre-task planning x online planning, $p = .180$). However, Ellis and Yuan (2004), Johnson et al. (2012), and Tabari (2016) reported contrasting findings. Tabari explained that one reason for the difference in the 2016 and 2017 studies was the use of different types of writing tasks—descriptive writing in Tabari (2016) and argumentative writing in Tabari (2017). He stated that the participants were able to benefit from pre-task planning more

when they engaged in the argumentative task because it is more cognitively demanding than the descriptive task. Due to learners' limited attentional resources, the impact of planning is more significant when they were engaged in complex tasks. Tabari's argument is partially supported because Ellis and Yuan (2004) also used a descriptive writing (story telling) task, but Johnson et al. used an opinion writing task, which is similar to argumentative writing. Although there are some inconsistent results, writing genre likely influence syntactic complexity.

Tabari (2022) updated his study using a mixed-method design and found that three pre-task groups performed significantly better than a no planning group regarding writing fluency and syntactic complexity. One hundred and eighty-five Iranian participants completed a narrative writing task in one of four conditions: pre-task planning, online planning, pre-task and online planning, and no planning. After the writing task, they participated in video-stimulated recall interviews. All three planning groups outperformed the no planning group regarding writing fluency and syntactic complexity ($p < .05$). Tabari concluded that Kellogg's writing model (1996), which states that pre-task planning can reduce the cognitive burden of the working memory required for formulation and promote written production, was supported. Another finding was that the comments from the video-stimulated recall interviews revealed that the pre-task planning condition was more effective for promoting the development of content and organization, while pressured online planning facilitated the translating and monitoring system.

The Effect of Individual and Collaborative Planning on Writing

While some researchers have focused on the effect of planning (e.g., with/without planning, with/without time pressure), other researchers have examined the effect of participation types—individual vs. collaborative (group or pair work)—on writing performance (Abrams & Byrd, 2016; Dobao, 2012; Doe & Figueroa, 2015; Tavakoli & Rezazadeh, 2014). For example, learners can plan individually, collaboratively in small groups, or with a teacher.

The common findings of Dobao (2012) and Tavakoli and Rezazadeh (2014) were that individual planning was more effective for improving writing fluency while collaborative planning enhanced syntactic accuracy. Dobao (2021) analyzed the effect of collaborative pre-writing tasks on L2 writing with 111 students studying Spanish in an American university. The participants were divided into three groups who performed a jigsaw task in individual, pair, and group conditions. After performing the pre-task in the assigned condition, each participant completed an individual writing task. The results showed that the students who worked alone produced longer texts than those in the other two groups. The students who worked individually performed more fluently than those who worked in pairs or groups because the latter groups had to engage in lengthy discussions whenever they made decisions; as a result, Dobao (2012) assumed that the students who worked alone were able to spend planning time more efficiently. On the other hand, working in small groups resulted in higher syntactic and lexical accuracy because collaborative work led to more occurrences of language-related episodes (LREs), which are defined “any part of a dialogue where the students talk

about the language they are producing, question their language use, or correct themselves or others” (Swain & Lapkin, 1998, p. 326). By working with their peers, they had opportunities to talk about and solve language problems, and eventually, they succeeded in solving many of the problems accurately.

Tavakoli and Rezazadeh (2014) also reported similar results in their examination of the effect of individual planning and collaborative planning on CAF. The researchers, who focused on the argumentative writing of 94 Iranian EFL learners, reported that collaborative planning enhanced syntactic accuracy (error free T-units: $F = 32.66$, $p < .01$; error free clauses: $F = 5.94$, $p = .02$), while individual planning promoted fluency (average words per text: $F = 4.22$, $p = .04$; average T-units per text: $F = 8.18$, $p = .01$; average clauses per text: $F = 10.64$, $p = .01$). Neither condition positively affected syntactic complexity (clauses per T-unit: $F = .78$, $p = .18$; dependent clauses percentage: $F = .20$, $p = .66$). As indicated by Dobao (2012), the participants who took part in collaborative planning had to spend time to integrate other people’s opinions and that could be one reason why the individual planning group outperformed the collaborative planning group in terms of writing fluency. However, the collaborative planning group had an advantage in that they produced more accurate texts, a result that supported previous studies (Lapkin & Swain, 2000; Lapkin et al., 2002; Swain, 1998). Tavakoli and Rezazadeh (2014) indicated that interaction can lead to more accurate grammatical performances because through the dialogues that occur during collaborative works, the participants had opportunities to talk about and solve language-related problems (Swain, 1998). These experiences raised their attention to linguistic form, and as a

result, it led to greater accuracy. Tavakoli and Rezazadeh (2014) also mentioned the social constructionist point of view of learning stated in Swain (1995) and proposed that “collaborative planning provides learners with chances for meaningful communication and involves them in cognitive processes which can be a source for L2 learning” (p. 99), suggesting that collaborative work has the potential to lead to the subsequent learning of L2 linguistic forms.

While Dobao (2012) and Tavakoli and Rezazadeh (2014) indicated that individual planning positively affects fluency and collaborative planning enhances accuracy, Doe and Figueroa (2015) and Abrams and Byrd (2016) reported different results. In Doe and Figueroa, 51 first-year university students received lists of questions related to the freewriting topic before writing. Half of them ($n = 26$), who were assigned to a collaborative planning condition (labelled *speaking planning* in the study), discussed the questions with their classmates while the other half ($n = 25$), who were in an individual planning condition (*writing planning*) condition, answered the questions individually and took notes on a piece of a paper. In terms of the total number of words produced, the writing planning group wrote more than the speaking planning group for each session (Freewrite 2: Speaking $M = 109.23$, $SD = 35.17$, 95% CI [93.23, 125.25]; Writing $M = 118.43$, $SD = 25.30$, 95% CI [108.99, 127.88]; Freewrite 6: Speaking $M = 127.62$, $SD = 41.93$, 95% CI [108.53, 146.70]; Writing $M = 150.23$, $SD = 38.96$, 95% CI [135.68, 164.78]; Freewrite 7: Speaking $M = 128.86$, $SD = 35.58$, 95% CI [112.66, 145.05]; Writing $M = 132.60$, $SD = 37.89$, 95% CI [118.45, 146.75]). However, there was no statistically significant difference between the collaborative and individual

conditions for syntactic accuracy, syntactic complexity, or writing fluency. The authors indicated that one possible reason to explain the lack of statistical significance was due to the short intervention period, but both types of planning conditions were effective because the descriptive statistics showed that the mean scores for the CAF measures for both groups improved after one semester. The results of these two studies suggested that participation types have different impacts on writing; however, more studies are needed before arriving at a firm conclusion.

Similarly, Abrams and Byrd (2016) did not find the superiority of collaborative planning on writing fluency or syntactic complexity but they did for accuracy ($t = 2.34$, $p = .03$, Cohen's $d = 0.95$), lexical richness ($t = 1.85$, $p = .05$, Cohen's $d = .76$), and overall writing scores ($t = 2.91$, $p < .01$, Cohen's $d = 1.19$). One group completed a meaning-focused pre-writing task collaboratively, while the other group engaged in a grammar-focused pre-writing task in pairs. The participants were 24 beginning level German learners. Contrary to previous studies, collaborative planning did not promote writing fluency. One possible reason for this result was whether the tasks were meaning-focused or grammar-focused, the learners could activate the relevant linguistic knowledge necessary to complete the writing task. As a result, the participants benefitted from both tasks, which could explain why the difference between the two tasks was insignificant.

Other researchers have investigated the effect of individual and collaborative writing in conjunction with task complexity and learners' language proficiency levels. To explore the effect of task complexity (less complex vs. more complex) and

proficiency (intermediate vs. advanced) on CAF, Ruiz-Funes (2015) analyzed the essays of 32 Spanish learners studying in an American university. For the intermediate level students, the less complex task was a personal essay, and the more complex task was an expository essay. For the advanced-proficiency students, the less complex task was an analytical essay and the more complex task was an argumentative essay. The complexity of these tasks was decided based on the demand of the cognitive processing required for information organization and information type (Skehan & Foster, 1999, 2001) and reasoning demand (Robinson, 2001). The descriptive statistics for both proficiency groups showed a similar trend; the more complex task promoted syntactic complexity while the less complex task led to enhanced syntactic accuracy and writing fluency, indicating that proficiency alone does not affect CAF measures. However, when subgroups were created based on the overall quality scores of the essays and each proficiency group was compared, a different trend was observed. In the case of the higher performers in the advanced-proficiency group, the more complex task resulted in the enhancement of all CAF measures but for the lower performers in the advanced-proficiency group, the CAF measures of the less complex task were higher. For the intermediate-proficiency group, three subgroups had similar patterns: higher scores for syntactic complexity with lower syntactic accuracy and writing fluency for the more complex task, and the opposite results for the less complex task. According to the authors, the more complex task had negatively affected the writing performance of the low-proficiency learners because they were able to focus on only one aspect of the language due to limited memory capacity in the L2. The more complex task added an

extra burden on their working memory and as a result, a trade-off effect between syntactic complexity and syntactic accuracy and syntactic complexity and writing fluency occurred. On the other hand, the language proficiency of the high performers among the advanced-proficiency learners was well-developed and their language processing was more automatized, which allowed them to focus their cognitive resources on all aspects of the language without experiencing a trade-off effect. Although there was no statistically significant differences among any of the measures due to the small sample size, Ruiz-Funes suggested that L2 proficiency can be a contributing factor that affects writing task performance in conjunction with the level of writing expertise.

Zhang (2022) also stated that the success of collaborative planning depends on the cognitive load of the task. She investigated the effect of task complexity on individual and collaborative writing with 69 EFL learners who were divided into two groups: individual and collaborative writing groups, and who completed two writing tasks. For a simple task, the participants wrote a letter to their friend about a traditional festival. They were provided with the task structure and background information about the topic. For the complex task, they wrote a letter about traditional art without any external resources. The texts were analyzed for syntactic complexity, syntactic accuracy, writing fluency, lexical diversity, and functional adequacy. The results revealed that the collaborative group significantly performed better than the individual group in terms of writing fluency in both simple and complex tasks (Simple: $Z = 2.37$, $p = .018$, $r = .34$; Complex: $Z = 3.01$, $p = .003$, $r = .51$). In addition, the collaborative

group outperformed the individual group regarding syntactic accuracy in the complex task (Errors per 100 words: $Z = 3.10$, $p = .002$, $r = .46$; Error-free clauses/total clauses: $Z = 2.63$, $p = .009$, $r = .41$; Error-free T-units/total T-units: $Z = 2.34$, $p = .009$, $r = .37$). However, the individual group scored higher than the collaborative group for functional adequacy in the simple task. In sum, the findings indicated that collaborative writing positively affects writing fluency and syntactic accuracy when task structure is provided. However, when task structure is unavailable, the individual writing group performed better than the collaborative writing group in terms of functional adequacy. The author suggested that the results can be explained by the trade-off effects of cognitive resources and indicated that whether the learners can benefit from collaborative writing depends on the cognitive load of the task.

Kang and Lee (2019) also explored the effect of task complexity (simple vs. complex) and a pre-task planning type (individual vs. collaborative) on syntactic complexity, syntactic accuracy, writing fluency, and lexical complexity. Similar to Ruiz-Funes (2015), and Zhang (2022), they found that collaborative planning enhanced writing fluency. Forty Korean EFL learners were divided into a simple task group and a complex task group. The participants in each group completed the pre-task planning in two conditions, one individually and the other in a collaborative setting. The writing was a narrative, story-telling task, and the simple task had only one main character, while the complex task had more than three characters. The findings indicated that the participants performed significantly better in collaborative planning than in individual planning in terms of writing fluency for both tasks, $F(1, 38) = 8.847$, $p = .005$, partial η^2

= .189. For lexical complexity, MTLTD was significantly higher in collaborative than individual planning but in the simple task group only, $F(1, 38) = 10.852, p = .002$, partial $\eta^2 = .222$. Although collaborative writing showed an advantage over individual writing for some syntactic complexity measures, the impact was not strong enough for statistical significance. Neither task type nor planning type significantly affected accuracy.

Kang and Lee indicated that the advantage of collaborative planning for writing fluency could be due to the rehearsal effect. Planning in pairs made it easier for the participants to access information. In addition, the authors stated that the Crossover Interaction Hypothesis (Kirschner et al., 2011), which proposes that collaborative learning is more advantageous for high-complexity tasks, could explain why individual planning functioned better than collaborative planning in the simple task for lexical complexity. For the simple task, the opportunity to interact was less fulfilling than in the complex task because the advantage of exchanging L2 knowledge did not pay off. As a result, the participants could use their time more efficiently in individual planning. This interpretation differs from Ruiz-Funes (2015) and Zhang (2022), who reported that collaborative planning with a complex task could negatively affect writing performance. Kang and Lee concluded that collaborative planning positively affects writing fluency and possibly syntactic complexity, but its effect is influenced by task complexity.

Chen and Lin (2021) proposed an advantage for collaborative learning, and although their study was not focused on L2 writing, it indicated how interaction can facilitate the understanding of metalinguistic and pragmatic knowledge. Thirty-three

Taiwanese EFL learners engaged in a multiple-choice discourse completion task individually or in pairs. The participants in both groups verbalized their thinking process as they completed the task. The recorded speeches were analyzed, focusing on the use of strategies for pragmatic-related episodes. The findings indicated that although the participation type did not significantly affect the multiple-choice discourse completion task scores, there was a significant difference between the two groups on strategy use. Furthermore, the results revealed that collaborative learning led to more negotiation and scaffolding. For example, intuitive deletion was used frequently by both groups. However, when working in pairs, the participants had to logically explain the reason for deleting a distractor to their partner. They suggested that the collaborative pragmatic task could enhance the understanding of metalinguistic knowledge and raise learners' consciousness of the target pragmatic features.

To summarize collaborative planning generally facilitates fluency and possibly syntactic and lexical diversity. However, the result is influenced by task complexity or factors such as the learners' language proficiency. Further research is needed to clarify these points.

Rehearsal

Pre-task planning, which takes place prior to task performance, can be divided into strategic planning and rehearsal (Ellis, 2005). While strategic planning allows learners to consider the content and language they use before writing, rehearsal gives them opportunities to practice the task before the main performance.

Rehearsal, which is also known as task repetition, provides learners with an opportunity to practice a task before the main performance (Ellis, 2005). Researchers have investigated the effect of rehearsal on oral performances and these studies have suggested that rehearsal positively affects learners' subsequent performances of the same task. For example, a study by Gass et al. (1999) looked at 103 Spanish learners whose L1 was English. One group watched the same video with no audio three times and orally described the events in Spanish. The same procedures were followed for the other group, but different videos were used each time. The researchers found that watching the videos of the same content positively affected holistic score ratings (sum of the differences between Test 1 and Test 3: same content = 4.56, different content = 3.39) the accuracy of the use of *estar* (the percentage of the participants who improved: same content = 40%, different content = 30%), and lexical complexity (percentage increase in content words: same content = 11.6%, different content = 5.6%). However, when a new video was used for Test 4, the different content group outperformed the same content group in terms of the holistic score and accuracy. The researchers, therefore, concluded that task repetition generally improves oral performance because it frees up learners' attention and thereby allows them to focus on linguistic aspects, but this benefit is not sustained when different content is used.

Bygate (1996, 2001) supported the findings reported by Gass et al. (1999). Bygate (1996) had one speaker repeat the same video narrative task twice on two occasions. When performing the task the second time, the participant was able to use more sophisticated lexis and grammatical structures compared with the first time. In a

larger study, Bygate (2001) compared the effect of specific task repetition—carrying out exactly the same task in a subsequent session—and task-type repetition—carrying out a new version of the same task type. He found that rehearsal enhanced oral fluency and syntactic complexity in the case of specific task repetition, whereas the effect of rehearsal was small in the case of task-type repetition. Based on these results, he concluded that task repetition is effective for improving oral performance because the participants were able to benefit from a priming effect that lasted even 10 weeks after the first session. When the same task was repeated, the learners could recycle ideas and output used in the previous task and more cognitive resources were available for language processing, but the effect of task type repetition was limited. Bygate stated that more task exposure would be necessary to gain stronger support for the effectiveness of task type repetition.

Few researchers have focused on the effect of task-repetition on writing performance. Nitta and Baba (2014), investigated the effect of two types of task repetition: specific task repetition and task type repetition. Unlike with oral tasks, the researchers found that the effect of specific task-type repetition was limited. The participants, 46 Japanese university students, completed ten-minute freewriting sessions over 30 weeks in which they wrote about the same topic for two consecutive weeks. To examine the effect of topic repetition, Nitta and Baba compared the texts written for the first time and those written one week later on the same topic focusing on five measures: the number of words produced as a measure of writing fluency, average sentence length and sentence syntax similarity as measures of syntactic complexity, and word frequency

values from the CELEX corpus and the measure of Textual Lexical Diversity (MTLD) as measures of lexical diversity. They found no statistical difference in the values of the five measures between the texts written in two consecutive weeks. For example, in the case of text length, the measure that was most influenced by task repetition, only five out of 14 sessions (35.7%) showed increases for one group and five out of 13 sessions (38.5%) for the other group. However, they found a stronger impact of repeating the same type of task on writing development. To examine the longitudinal effect of task type repetition, the compositions written in Week 1 and Week 30 were compared. Two syntactic complexity measures showed significant increases in both classes and MTLD increased in one class: Average sentence length: $F = 19.02, p < .001$; Sentence syntax similarity: $F = 19.98, p < .001$; MTLD: $F = 17.97, p < .001$. The researchers concluded that in the case of writing, the benefit of repeating the same task is limited compared to oral tasks, but the effect of repeating the same type of task over a long period of time positively affects the development of syntactic complexity and lexical diversity.

The effect of repeating the same type of task (i.e., task procedural repetition) and repeating the same task on syntactic complexity, syntactic accuracy, and writing fluency was also examined by Amiryousefi (2016). The participants, 70 low-intermediate proficiency Iranian EFL students, completed story-telling tasks while looking at eight pictures. They were randomly assigned to one of two groups: (a) task procedural repetition, a group that repeated the same type of task but was given different content each time, and (b) task repetition, a group that repeated the same task five consecutive times over five weeks. Seven CAF measures, the number of words, T-units, and clauses

per composition were calculated as measures of writing fluency, the percentage of dependent clauses to total clauses and the percentage of dependent clauses to T-units were measures of syntactic complexity, and the percentage of error free clauses and error free T-units were the measures of syntactic accuracy. The pretest and posttest results were compared using independent-sample *t*-tests and Mann-Whitney *U* tests. The results of a within-group comparison indicated that both groups performed better on the posttest than the pretest, but the group that repeated the task five times scored significantly better on the posttest for all three fluency measures (number of words: $t[34] = 277.34, p < .001$; number of T-units: $t[34] = 2.473, p = .20$; number of clauses: $t[34] = 8.251, p < .001$), one accuracy measure (the percentage of error-free clauses: $t[34] = 2.792, p = .009$), and only two fluency measures (the number of words: $t[34] = 90.095, p = .000$; clauses per text $t[34] = 2.980, p = .005$) improved for the procedural repetition group. Moreover, the results of a between-group comparison for the posttest showed that the task repetition group significantly outperformed the procedural repetition group on all three fluency measures and the percentage of error-free clauses (number of words: $t[68] = 13.344, p < .001$; number of T-units: $t[68] = 2.595, p = .012$; number of clauses: $t[68] = 6.040, p < .001$) and one accuracy measure (percentage of error-free clauses: $t[68] = 2.251, p = .028$). Amiryousefi suggested that the results supported the limited processing capacity of L2 learners. When exactly the same task was repeated, the cognitive burden was reduced for the participants in the task repetition group because they could reuse the content and language they had used in the previous session. As a result, more attention was available to focus on linguistic aspects such as

fluency or syntactic accuracy. Amiryousefi concluded that task repetition increases the learners' capacity for language processing and helps them to produce more sophisticated language.

The results of Nitta and Baba (2014) and Amiryousefi (2016) imply that both task repetition and procedural repetition positively affect writing performance if they are applied over the long term. When repeated enough times, repeating exactly the same task might be more effective than procedural repetition, but unlike oral performances, repeating it once a week does not result in improvements in the case of writing. However, there are not enough investigations of the effect of task repetition on writing to arrive at firm conclusions.

Reading for Writing

Another way to categorize planning concerns unguided planning in which learners plan freely or guided planning in which learners are advised to focus their attention on particular linguistic forms, meaning, or both (Ellis, 2005). Teacher-led planning is one example of guided planning. Compared with the studies that were focused on L2 learners' speech (Foster & Skehan, 1999; Kawauchi, 2005; Mochizuki & Ortega, 2008; Ogawa, 2016), there have been investigations on the effect of teacher-led planning focusing on L2 writers.

Kawauchi (2005), who focused on oral speech, reported that students provided with model passages borrowed lexical items from the models. She explored how three planning conditions—writing (writing a draft), rehearsing (saying what they want to say aloud), and reading (reading model L2 input)—and English proficiency level (Low

EFL, High EFL, Advanced ESL) affected syntactic complexity, accuracy, and fluency on a narrative story-telling task. When planned and unplanned conditions were compared, performances in the planned condition were significantly better than the unplanned conditions in three out of seven syntactic complexity, syntactic accuracy, and fluency measures, but when CAF measures of the three planning conditions were compared, there was no significant differences among any of the groups. Although the planning conditions did not affect any of the CAF measures significantly, the author reported that qualitative data indicated that the students engaged in different activities for each of the planning conditions. For example, the use of some low-frequency words that appeared in the model passages were observed in the speech transcriptions of the reading group. She also noted that after engaging in the reading task, the participants focused on solving particular linguistic problems whereas the writing and rehearsal tasks drew their attention to non-linguistic aspects such as adding their own interpretations or comments to the story.

Ogawa (2016), in another study focused on oral speech, investigated the effect of teacher-led planning and three other planning conditions: solo written brainstorming, paired-interactive planning, and no planning on syntactic complexity, syntactic accuracy, and fluency on a monologue task. In the teacher-led planning group, the students received handouts in which sample passages were printed, and they silently followed the passages and listened while the teacher read the passage aloud. Ogawa reported that teacher-led and interactive planning had a significant effect on improving complexity: Clauses per AS-units, $F[3, 228] = 3.24, p = .02$; Mean length AS-units,

$F[3, 228] = 4.12, p = .007$. Citing evidence that the students in the teacher-led planning group reproduced the ideas and used the lexical items provided by the model passage, she reported that those students formed more syntactically complex sentences because the teacher's model allowed them to allocate more attentional resources to the formulation and articulation stages by easing cognitive processing in the conceptualization stage.

Reading and writing are connected in classroom settings because students are often given reading assignments before writing. According to Hirvela (2016), there are two aims of reading for writing. The first is to provide models for students to learn about writing. By referring to model essays, students can visualize the form their essays should take in the target language. In addition, reading texts provides students with useful target language input. Stotsky (1995) stated, "reading experience would seem to be the chief source of a developing writer's syntactic, generic, and lexical knowledge" (p. 73). Learners can scaffold lexis and syntax in the reading text and apply them in their own writing.

The second aim is to provide information about the topic so that learners have ideas to write about. A number of L1 researchers have shown that topic familiarity positively affects the quality of writing performances (Bereiter & Scardamalia, 1987; Calkins, 1986; McCutchen, 1986). For example, McCutchen (1986) asked students to write compositions about football, and found that football experts wrote more coherent texts and elaborated on ideas more than non-experts. Similar results were reported by Bereiter and Scardamalia (1987) who found that when children wrote compositions

about familiar topics, they produced more content during planning compared with unfamiliar topics. Furthermore, McCutchen et al. (1997) asked their participants to revise their essays on both familiar and unfamiliar topics. The researchers found that topic familiarity did not influence the correction of spelling errors, but the participants corrected more meaning-related errors when writing about the familiar topic.

McCutchen (2000) stated that these results can be explained by the notion of long-term working memory proposed by Ericsson and Kintsch (1995). Unlike short-term working memory, which has a limited capacity, long-term working memory has a retrieval function and connects an item activated in short-term working memory with related items in long-term memory, allowing writers to access the extensive conceptual resources stored in long-term memory. For correcting spelling errors, topic knowledge was not necessary, but for correcting logical problems, they needed to access resources in long-term working memory. Thus, topic knowledge helped the participants use their limited cognitive resource more efficiently.

Studies targeting English learners also indicate the positive effect of topic familiarity on writing product. Esmaeili (2002) investigated the effect of a pre-reading passage on the scores of writing and reading recall tasks. The participants were 34 intermediate-proficiency adult ESL students who completed the writing tasks in two conditions: (a) a passage thematically related to the writing task and (b) a passage thematically unrelated. The results revealed that when the pre-reading passage was thematically related to writing, the holistic scores of the writing task, $F(1, 33) = 134.28$, $p < .001$, and the recall task $F(1, 33) = 134.28$, $p < .001$, were significantly higher than

when the passage was unrelated to the theme of writing. The results indicated that the relevance of the themes between reading and writing facilitated writing performance.

Another study on topic familiarity targeting English learners was conducted by Kessler et al. (2012). To investigate the effect of topic familiarity on subsequent writing performance, the researchers examined the essays written by 60 ESL learners on two topics: a familiar topic (use of a cell phone while driving) and an unfamiliar topic (safety of electronic cigarettes). They found that topic familiarity had a significant effect on two out of 12 complexity measures, two out of two lexical complexity measures, and two out of four accuracy measures. Topic familiarity had no significant effect on fluency. The familiar topic resulted in significantly higher complexity than the unfamiliar topic, Clauses per T-unit: $p < .01$, $\eta_p^2 = .061$, $d = .48$); Verb phrases per T-unit: $p < .001$, $\eta_p^2 = .186$, $d = .90$, although for accuracy, the unfamiliar topic resulted in significantly higher accuracy than the familiar topic: Error-free T-unit ratio: $p < .001$, $\eta_p^2 = .171$, $d = -.90$; Morphological errors/100 words: $p < .001$, $\eta_p^2 = .096$, $d = .64$. For lexical complexity, the unfamiliar topic led to significantly higher AWL while for MTLT, the score of the familiar topic was better than that for the unfamiliar topic: Average word length: $p < .001$, $\eta_p^2 = .38$, $d = -1.52$; MTLT [Measure of Textual, Lexical Diversity] $p < .001$, $\eta_p^2 = .132$, $d = .73$. They indicated that the findings could be explained by Skehan's (1998) Trade-Off Hypothesis. The task on a familiar topic enhanced complexity because the cognitive demand was low, pushing learners to use more complex language. On the other hand, the cognitive load was high for the unfamiliar topic, and they used simpler language, resulting in higher accuracy.

Factors That Affect Writing Performance

Some researchers have examined what constitutes the reading-to-write construct (Delaney, 2008; Kim & Pae, 2021; Yang & Plakans, 2012). Delaney (2008) investigated what factors constitute the reading-to-write construct with 139 participants—50 native speakers, 62 ESL learners, and 27 EFL learners—who completed two reading-to-write tasks: a summary and response essays about the same pre-reading passage. In the summary essay, the participants summarized the text's key ideas, while in the response essay, they responded to the text's main ideas. The finding was no significant correlation between writing ability and either writing task. In contrast, reading ability significantly correlated with the summary task ($r = .28, p < .01$) and the response task ($r = .38, p < .01$). The author suggested that reading-to-write ability is different from writing ability that requires no reading because the participants had to identify the key ideas in the text and condense them to write a summary. For the response essay, they had to comprehend the text and decide which ideas to respond to while reflecting on their own opinions. Delaney concluded that the reading-to-write construct is weakly correlated to reading ability but could be independent of writing ability. The author found that language proficiency and educational level also have a modest effect on the reading-to-write construct.

Although the correlation of writing ability and writing task was weak in Delaney's (2008) study, Kim and Pae (2021) reported that L1 writing was one of the factors that affected a writing task. Using a structural equation model, Kim and Pae (2021) explored the effect of six linguistic and affective factors on L2 writing. The six

variables were L1 and L2 writing ability (narrative and argumentative writing in both languages), L2 reading comprehension (literal and inferential), L2 proficiency (knowledge of English grammar and vocabulary), L2 motivation, and L2 anxiety. The participants, 298 Korean 10th graders, completed two writing tasks: narrative writing (less cognitively demanding) and argumentative writing (more cognitively demanding). The results indicated that L1 writing, L2 proficiency, and L2 writing anxiety had a direct effect on both L2 narrative and L2 argumentative writing (narrative: L1 writing, $\beta = 0.45, p < .05$; L2 proficiency, $\beta = 0.31, p < .05$; anxiety, $\beta = -0.26, p < .05$; argumentative: L1 writing, $\beta = 0.48, p < .05$, L2 proficiency, $\beta = 0.39, p < .05$ anxiety, $\beta = -0.31, p < .05$). On the other hand, L2 reading had a direct impact on only the argumentative task ($\beta = 0.22, p < .05$). Based on this finding, the authors indicated that reading contributes more effectively to argumentative writing than narrative writing. They suggested that through inferential reading activities, learners can develop inferential skills and learn how to deliver their message logically and persuasively.

For integrated reading-listening and writing tasks, comprehension and production abilities are not the only factors that affect writing performance. Yang and Plakans (2012) conducted a study on the construct of an integrated reading-listening and writing task with 161 ESL students studying at a university in the United States. They analyzed the relationship between integrated writing task performance and strategy use using a structural equation model. The findings indicated that the construct comprised three categories of strategy use: self-regulatory (i.e., monitoring), discourse synthesis (i.e., organizing, selecting), and test-wiseness strategies (i.e., copying, using writing

models, patchwriting). They reported that self-regulatory strategy use had a positive direct effect on discourse synthesis strategy use ($\beta = 0.77$) and test-wisness strategy use ($\beta = 0.41$). Discourse synthesis use had a positive direct effect on integrated writing performance ($\beta = 0.33$) and test-wisness strategy use had a negative direct effect on the performance ($\beta = -0.11$). The author concluded that the construct of the integrated task is multifaceted, involving more than just comprehension and writing abilities.

These studies have shown that integrated writing is affected by factors such as L1 and L2 writing and reading ability, writing genre, educational background, L2 writing anxiety, and the ability to use self-regulatory or discourse synthesis strategies.

Meaning-Focused Writing and Writing Development

In English-as-a-foreign-language (EFL) contexts where the opportunity to produce target language output is limited, extensive writing activities are necessary to develop writing fluency. Timed writing, also known as freewriting or speedwriting, is a meaning-focused activity used to develop writing fluency. In general, students have 10 to 15 minutes to write as much as they can without worrying about errors. There are various ways to implement timed-writing in class in terms of time, content, and review. For example, in some cases, students have complete freedom in selecting writing topics while in other cases, instructors provide the topics or themes (Tanner, 2016). In this study, I focus on timed-writing in which the topics are fixed and students express opinions about those topics.

Introducing freewriting in the writing classroom provides several advantages (Polio, 2012). One advantage is that freewriting allows students to produce the foreign

language. Another advantage is that it provides them with opportunities to practice language they have acquired previously, which is especially important for those studying in EFL environments, where opportunities for language production outside the classroom are limited. In addition, freewriting allows students to take risks when writing. Liao and Wong (2010) reported that freewriting effectively provided an “anxiety-free context” (p. 141) to Taiwanese English learners, who were often reluctant to take risks in writing because they often write for examinations that are graded based on grammatical accuracy. An additional advantage of freewriting is that its features meet the four criteria that, according to Nation (1991, 2013), a fluency activity must meet. Nation (1991) suggested that to develop fluency: (a) the language used in the activity is known to the learners; (b) the activity places a degree of pressure on the learners to process the target language faster than usual; (c) the activity requires the learners to use large quantities of the target language, and; (d) the activity is meaning-focused. Freewriting is a useful activity for improving writing fluency because it meets all four of the above criteria.

Some researchers have questioned the effectiveness of freewriting because it is unstructured and teachers do not correct errors. Tanner (2016) stated that freewriting “...is not a panacea” (p. 10); however, students can benefit from freewriting where linguistic variables are concerned if teachers provide feedback on grammar or vocabulary. Baba and Nitta (2010), Nitta and Baba (2014), and Nguyen (2015) provided evidence that freewriting is beneficial even without corrective feedback. Baba and Nitta (2010) investigated the effect of student engagement in 10-minute freewriting tasks for

30 weeks on writing fluency, syntactic complexity, and lexical diversity. The participants were 46 Japanese English learners in two university English classes, and the compositions written in Week 1 and Week 30 were compared. The researchers found that two syntactic complexity measures, average sentence length ($F = 19.02, p < .001$) and sentence syntax similarity ($F = 19.98, p < .001$), increased significantly in both classes, and one lexical measure, MTLN ($F = 17.97, p < .001$), increased in one of the two classes. They stated that one reason for why the freewriting task was the most effective for the development of syntactic complexity was that one-year of extensive writing practice raised the learners' awareness of formal aspects of writing. After each writing session, the students had opportunities to see their own writing and they attempted to write better the next time. The researchers also proposed the existence of competition between fluency and syntactic complexity/lexis. According to the authors, less proficient learners tend to focus on improving fluency at the expense of syntactic complexity and lexical diversity, but as they become more proficient through task repetition, their focus shifts to syntactic or lexical aspects, and they downplay the improvement of fluency.

To investigate the effect of a freewriting task on writing development, Nguyen (2015) conducted a study with 110 Vietnamese university students. A positive effect was found on writing development after the students engaged in freewriting tasks for ten weeks. The students in the intervention groups, two separate classes with the same intervention, completed a seven-minute freewriting activity three times a week for ten weeks, while the comparison group only completed the tasks in the first and last classes.

Six CAF measures were compared between the intervention and comparison group: words per seven minutes, number of sentences, and words per sentence for fluency; grammatical mistakes per 100 words and overall quality (global judgement on a five-point scale) for accuracy; and the percentage of the complex sentences in the total number of sentences for syntactic complexity. Nguyen found that one accuracy measure, overall quality ($F = 35.227, p < .001$) and one fluency measure, total number of words ($F = 40.388, p < .001$) were significantly higher for one of the two intervention groups compared with the comparison group. The author reported that constantly engaging in a seven-minute writing activity was helpful for developing fluency because English learners in EFL environments do not spend enough time writing in English and the activity gave them extra opportunities to practice.

While the studies by Nitta and Baba (2014) and Nguyen (2015) were focused on the effect of meaning-focused writing activities on syntactic complexity, syntactic accuracy, and writing fluency development in foreign language classes in EFL environments, Bulté and Housen (2014) and Mazgutova and Kormos (2015) focused on the writing development of L2 students in English speaking environments. Bulté and Housen (2014) compared the development of syntactic and lexical complexity of adult ESL students before and after completing a four-month intensive academic writing course. Significant differences between the pretests and posttests were found for seven out of ten syntactic complexity measures (Mean length of sentence: $t = -2.956, p = .005, d = 0.441$; Mean length of T-unit: $t = -3.140, p = .003, d = 0.468$; Simple sentence ratio: $t = 2.887, p = .006, d = 0.441$; Compound sentence ratio: $t = -2.673, p = .012, d =$

0.393; Coordinated clause ratio: $t = -2.102$, $p = .041$, $d = 0.313$; Mean length of finite clause: $t = -3.310$, $p = .002$, $d = 0.493$; Mean length of noun phrase: $t = -2.726$, $p = .009$, $d = 0.406$) but no significant differences for lexical measures (G, the variation in and the number of word types used: $t = 0.153$, $p = .879$, $d = 0.023$; D, the diversity in the use of word types): $t = 0.899$, $p = .374$, $d = 0.134$; AG, the variation in and the number of basic and advanced word type used): $t = -0.627$, $p = .534$, $d = 0.094$. Their study showed that syntactic and lexical complexity do not develop in parallel; they constitute separate dimensions of L2 performance and L2 proficiency at least for L2 learners. They also suggested that in the short or medium term, complexity, accuracy, and fluency do not develop linearly but rather fluctuate. For this reason, longitudinal studies are necessary to fully capture the trajectory of L2 development.

The findings reported by Mazgutova and Kormos (2015), however, indicated somewhat different results in terms of the development of lexical complexity. Their participants were L2 students from two proficiency groups. One group consisted of upper-intermediate (CEFR C1 level) students who were planning to enroll in a postgraduate program, while the intermediate group consisted of students preparing for undergraduate university studies in the United Kingdom (CEFR B2 level). Unlike the participants in Bluté and Housen (2014), more growth was observed for lexical complexity than for syntactic complexity by the participants in Mazgutova and Kormos's study. After completing a one-month program, the scores of two out of six lexical complexity measures increased for the upper-intermediate group (squared verb variation: $Z = -3.123$, $p < .002$, $r = -0.44$; academic word list: $Z = -2.222$, $p < .026$, $r = -$

0.31) and all five measures for the intermediate group (MTLD: $Z = -3.296, p < .001, r = -0.62$; squared verb variation: $Z = -2.731, p < .006, r = -.52$; log frequency of content words: $Z = 2.166, p < .03, r = .41$; academic word list: $Z = -2.104, p < .035, r = -.40$); latent semantic analyses: $Z = -2.04, p < .041, r = -.39$). In the case of syntactic complexity, only one out of five measures was significant for the upper-intermediate group (syntactic structure similarity: ($Z = -3.287, p < .001, r = -.46$) and three out of five for the intermediate group (complex nominals ($Z = -2.045, p < .041, r = -.39$); modifiers per noun phrase: $Z = -2.542, p < .011, r = -.48$); and syntactic structure similarity ($Z = -2.355, p < .019, r = -.45$). In the discussion section, the authors reported that a one-month extensive English program in the United Kingdom provided the students with opportunities to learn new words incidentally both inside and outside the classroom. The researchers also found that the syntactic similarity posttest score was significantly higher than the pretest for both proficiency groups. Their hypothesis was that L2 learners first attempt to use a variety of syntactic forms, but when they reach a certain point, they tend to use similar syntactic constructions to maintain readability. If correct, that explains why syntactic complexity did not develop significantly compared to lexis.

To summarize the results of the above studies, syntactic complexity improved in Nitta and Baba (2014), Nguyen (2015), and Bluté and Housen (2014), but not as much development was observed for the more proficient learners in Mazgutova and Kormos's study. A significant change in lexical complexity was observed in Mazgutova and Kormos's study, but not in Bluté and Housen's study nor in the lower-proficiency

group in Nitta and Baba's study. These differences in the previous studies indicate that L2 syntactic and lexical complexity develop in complex ways and they are influenced by learners' backgrounds such as their L1 or L2 proficiency (Ortega, 2015). Because there are not enough longitudinal studies on writing development targeting Japanese English learners, more studies in this area are needed.

Gaps in the Literature

The effect of pre-tasks on writing has attracted the interest of many researchers, and a large number of studies with various types of tasks and conditions have been reported. However, the majority have concerned planning time or the effectiveness of strategic vs. online planning (Dobao, 2012; Doe & Figueroa, 2015; Ellis, & Yuan, 2004; Johnson et al., 2012) even though many other types of tasks are available. Regarding the effect of rehearsal or task-repetition, this study is unique because it is focused on the effect of pre-task planning in combination with oral rehearsal on writing performance. This issue is worth investigating because different results can be expected for different modalities. In addition, there has been little attention on the effect of reading a passage on L2 writing performance. Studies involving L1 writers have indicated that genre or topic familiarity positively affects writing quality (Bereiter & Scardamalia, 1987; Calkins, 1986; McCutchen, 1986), and one investigation concerned the effect of a reading passage on L2 oral performance (Kawauchi, 2005), but there has been only a small number of studies on the effect of reading a sample passage on L2 learners' writing performance.

Moreover, it is not clear what learners do during pre-task planning, alone or together with oral rehearsal or when reading and how the cognitive processes that take place during pre-writing activities are incorporated in the final written product. As mentioned by Rostamian et al. (2018), most previous studies have been focused on either cognitive processes or the quality of the final product. Little is known about “what learners allocate their limited attentional and memory resources to, how task manipulation affects the cognitive processing of writers, and how this manner of resource allocation may impact the resulting text in terms of writing quality” (Ong & Zhang, 2010, p. 218), and it is necessary to clarify how writers’ thinking and behavior during pre-task activities affect their writing performances.

Finally, numerous researchers have explored the effect of meaning-focused writing activities on L2 students’ writing development, but the data collection period was one semester in the majority of the studies. More longitudinal studies that last longer than one semester on meaning-focused writing activities and their effect on various aspects of writing development—syntactic complexity, syntactic accuracy, lexical diversity, and writing fluency—are needed. As indicated by Nitta and Baba (2014), writing development is a complex system and understanding the longitudinal effects of a task is as important as investigating the immediate outcome of the task if researchers are to fully understand L2 writing development.

Purposes of the Study

The first purpose of this study was to investigate how three types of pre-tasks—outlining, oral rehearsal, and reading—affect syntactic complexity, fluency, lexical

diversity, content, and comprehensibility of timed-writing texts. As mentioned above, research to date has shown that planning affects writing performance positively (Ellis & Yuan, 2004; Kellogg, 1996; Rostamian et al., 2018). However, the effect of oral rehearsal and reading a passage on L2 learners' writing, especially low-proficiency English learners, is unclear. Therefore, this study aims to investigate how different pre-tasks affect the writing of low-proficiency English learners.

The second purpose was to examine how the participants make use of each planning opportunity and how their cognitive processing during the pre-tasks influence the final written products. Although the need to examine what L2 writers actually do during the pre-task and on-line planning has been mentioned (Ellis & Yuan, 2004), the majority of the studies on writing pre-tasks have relied on quantitative data (Johnson et al., 2012; Ong & Zhang, 2010; Ong & Zhang, 2013; Tabari, 2016, 2017). Rostamian et al. (2018) is the only mixed-method study in which interview data were used to focus on the writing process. More such studies are needed. By conducting stimulated recall interviews, I explored how the participants benefited from each type of pre-task activity when producing their timed-writing texts. This process helped clarify how pre-task features affected the allocation of their limited cognitive resource while writing and how the allocation of their cognitive resources influenced the final written product.

The third purpose was to investigate how the repetition of meaning-focused writing tasks over one academic year contribute to L2 students' writing development. Some researchers have conducted studies on the effect of timed-writing on writing development, but the data collection period was just one semester (Bluté & Housen,

2014; Doe & Figueroa, 2015; Nguyen, 2015). To investigate the longitudinal effect of timed-writing activities on writing development, the current study was conducted for a full academic year with the same participants. In addition, this study adds new insight to the work of Nitta and Baba (2014), who conducted a one-year study but with no comparison group. By comparing the compositions written by the experimental group and comparison group at three time points, I aimed to examine the longitudinal effect of meaning-focused writing on the participants' linguistic development.

Research Questions

This study was guided by the following three research questions.

1. What is the effect of three planning conditions—outlining, oral rehearsal, and reading—on syntactic complexity, writing fluency, lexical diversity, content and comprehensibility?
2. How do the participants make use of each pre-writing task opportunity, and how is it reflected in the quality of their written products?
3. How does the experience of repeating the timed-writing task for one academic year contribute to the development of the participants' syntactic complexity, writing fluency, and lexical diversity?

CHAPTER 3

METHODS

In this chapter, I first outline the context of the study, including a general outline of the institution, its overall curriculum and selectivity, and the organization of the academic year. Next, I outline the content of the classes in the program relevant to the study and discuss how these classes are generally taught. Following this, I provide an overview of the participants in the study, who were drawn from six intact classes. This section includes a discussion of the experimental and comparison groups and information about the background of the teachers for the classes. Next, I discuss the research design for the study, with consideration of the quantitative and qualitative data collection. Following this, I describe the instruments, their development, the procedures used to collect quantitative and qualitative data, and the steps taken in the analysis of the data.

Context of the Study

This study was conducted at a private co-educational university located near Tokyo, Japan. Established in 1987, the university consists of three foreign language departments—English, Asian Languages, and Spanish and Portuguese—and a department of International Communication. According to the data of Japanese university rankings on *Pass-navi*, reported by Obunsha (2020), the *hensachi* (standardized rank score) of the school, where 50 indicates the 50th percentile, is between 45 and 50, meaning that it is ranked slightly below average. According to Japan Statistical Society (2012), *hensachi* is calculated as follows.

$$hensachi = \frac{10(x_i - \mu_x)}{\sigma_x} + 50$$

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2}$$

where n = number of scores in the sample, x_1 = each value in the data set, μ_x = mean, and σ_x = standard deviation

Each department has a different English curriculum. Some departments provide all-English courses taught by non-Japanese instructors. Other departments offer English courses taught by both Japanese instructors in Japanese and courses taught only in English by non-Japanese instructors. The students who participated in this study were double major students majoring in English and a second foreign language. In their first year, they take two English courses. One class is focused on reading, writing, and grammar and is taught by Japanese teachers, mainly in Japanese. In this class, students read textbook passages, strengthen the foundation of their grammar knowledge, learn English paragraph structure, and write paragraphs or short essays. The other class was an integrated four-skills course in which improving speaking and listening skills is emphasized. The students engaged in various communicative activities and gave presentations in English using iPad devices. The instructors were mainly native speakers of English, and the course is taught in English. Each class meets twice a week,

giving students four 90-minute English classes per week throughout the 30-week academic year. In their sophomore year, the students took a four-skills integrated English course that met three times a week, and in their junior year, they took a content-based English course that met twice a week throughout the 30-week academic year, all taught only in English. I submitted the form required by the institution and got approval from the university president and the department dean for conducting the study and collecting data.

Participants

The participants were 84 first-year students attending six mandatory first-year English classes. Of the eight English classes in the cohort, six conveniently selected classes participated in the study. The participants in three classes taught by me were assigned to an experimental group ($n = 42$), and the participants in the remaining three classes taught by another Japanese instructor were assigned to a comparison group ($n = 42$). Initially, 89 students were enrolled in the six classes; however, two students in the experimental group dropped out of the course, and three students in the comparison group did not sign consent forms. Thus, the data of these students were not included in the analyses.

The experimental and comparison groups differed in terms of their English proficiency. The students were placed in English classes based on the TOEIC scores they took just before entering the university. The three classes assigned to the experimental treatments were in the lower half of the English proficiency classes in the department (Classes A, B, and C), and the other three classes assigned to the

comparison group were in the upper half of the English proficiency classes (Classes D, E, and F). The average TOEIC scores by class and group are indicated in Tables 1 and 2. Several students in class D had studied abroad; however, most of the participants in both the experimental and comparison groups were from ordinary Japanese high schools.

Table 1. *Descriptive Statistics for the TOEIC Scores of Participants by Class in Study*

Class	Experimental group (<i>n</i> = 42)			Comparison group (<i>n</i> = 42)		
	A (<i>n</i> = 15)	B (<i>n</i> = 12)	C (<i>n</i> = 15)	D (<i>n</i> = 15)	E (<i>n</i> = 14)	F (<i>n</i> = 13)
<i>M</i>	390.71	388.21	350.67	527.50	422.33	451.07
<i>SD</i>	60.82	99.60	89.00	126.89	104.04	130.77
Maximum	510.00	535.00	485.00	740.00	610.00	640.00
Minimum	290.00	140.00	185.00	345.00	255.00	190.00

Table 2. *Descriptive Statistics for the TOEIC Scores by Group*

	Experimental group	Comparison group
<i>M</i>	375.93	468.67
<i>SD</i>	84.90	126.79
Maximum	535.00	740.00
Minimum	140.00	190.00

Note. Both the experimental group and comparison group consisted of three classes. Each of the three classes in the experimental group was assigned a different treatment every three weeks to counterbalance the effect of learner variables.

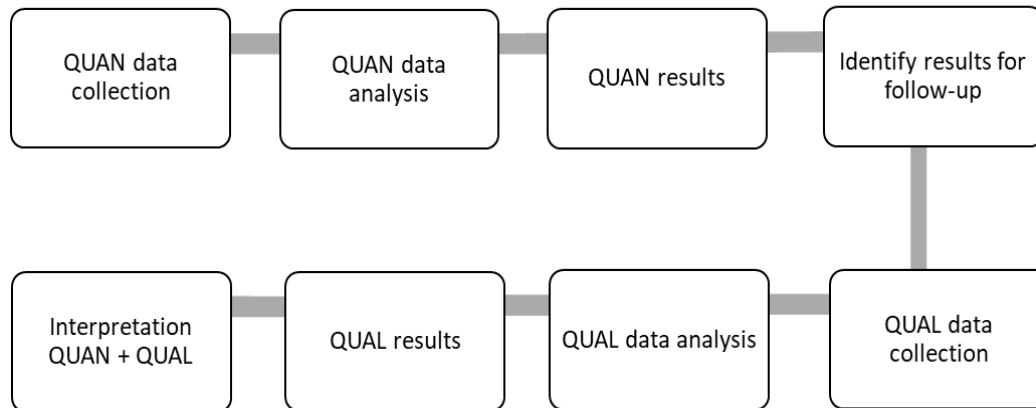
Research Design

A mixed-method explanatory design was employed in the present study. This design was adopted because although the quantitative data indicates the outcome of the intervention, it alone does not provide sufficient information regarding the cause of the results. Creswell and Plano Clark (2007) stated that the explanatory design, which is also called an explanatory sequential design, consists of two phases, an initial quantitative phase and subsequent qualitative phases.

Figure 2 shows the design of the study. In the first stage, quantitative data were collected and analyzed. The intervention period was nine months, and the texts were analyzed using text analysis software Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (Kyle, 2023) and Tool for the Automatic Analysis of Lexical Diversity (Kyle et al., 2023). In the second stage, three types of qualitative data were collected to supplement the results of the quantitative analysis. One was the data from interviews. The interview questions were designed based on the results of the first stage to explore the participants' cognitive process when engaging in the pre-writing tasks and timed-writing tasks. Another one was a post-course questionnaire that was conducted at the end of the academic year after the participants had completed all of the timed-writing sessions. It was conducted to explore how the participants perceived each of the three pre-tasks. The third one is the supplemental analysis of lexis. The similarity rate of the lexis between the reading passage and the timed-writing texts was compared to examine to what extent the passage had an influence on the lexical diversity in the timed-writing.

Creswell and Plano Clark (2007) explained that the explanatory model is suitable when researchers are interested in exploring “specific quantitative findings that need additional explanation” (p. 72) because qualitative data can be used to help explain or build on the results of the initial quantitative stage. One purpose of this study was to explore what the participants did while engaging in the three types of pre-writing tasks and reveal the relationship between the mental processes involved in each pre-task and the final written products. This design was suitable for this purpose.

Figure 2. *The Design of the Study*



QUAN = quantitative; QUAL = qualitative

From *Designing and Conducting Mixed Methods Research* (p.63), by J. W. Cresswell and V. L. Plano Clark, 2007, Sage. Adapted with permission.

Instrumentation

Timed-Writing Task

Timed-writing topics were selected carefully because topic familiarity can influence writing performance (McCutchen, 2000). It was important to select topics that did not require specialized knowledge. By adapting the writing topics used on previous Eiken tests and the TOEFL Writing Test (TWE) as well as original topics created by myself, 22 topics for opinion paragraphs, including four topics for the tests, were prepared and randomly assigned for each session. The topics used in this study are shown in Table 3.

Table 3. *Timed-Writing Topics*

Session	Topic
Practice	Being a university student is better than being a high school student (大学生でいることは高校生でいることよりも良い)
Pretest	(Participants were randomly assigned one of the topics) Do you agree or disagree? A. Traveling in a foreign country better than traveling in Japan. (海外旅行は国内旅行よりも良い) B. Living in a city is better than living in countryside. (都市に住むのは田舎に住むよりも良い) C: Having siblings is better than being an only child. (兄弟がいることは一人っ子よりも良い) D: Japanese style breakfast is better than western style breakfast. (朝食は日本食が洋食よりも良い)
1	Do you agree or disagree? Watching a movie in a theater is better than watching it at home. (映画は映画館で見る方が家で見るよりも良い)
2	Hokkaido is a better place to visit for a vacation than Okinawa. (休暇で訪れるなら沖縄より北海道が良い)
3	SNS is a good communication tool. (SNS はよいコミュニケーションの道具である)
4	It is a good idea to have pets. (ペットを飼うのは良いことだ)
5	It is good for high school students to have school uniforms. (高校生にとって制服があるのは良いことだ)
6	It is better for university students to live with their parents than to live on their own. (大学生はひとり暮らしより実家暮らしの方が良い)
7	Paying by electronic money (e.g. Paypay, LINE pay etc.) is better than to paying by cash. (電子マネー [e.g. Paypay, LINE pay など] は現金よりも良い)
8	Traveling by train is better than airplane. (移動手段は飛行機よりも電車が良い)
9	It is good for elementary school children to have smart phones. (小学生がスマホを持つのは良いことだ)
Posttest 1	(The topics are the same as Test 1)
10	Living in a house with a garden is better than living in an apartment. (マンションより庭つきの一軒家に住む方が良い)
11	iPhones are better than Android smartphones. (アンドロイドスマートフォンより、iPhone が良い)

Table 3 (continues).

Table 3 (continued).

Session	Topic
12	Traveling with friends is better than traveling alone. (友人と旅行するのは1人で旅行するより良い)
13	It is good for university students to do a part-time job. (大学生がアルバイトをするのは良いことだ)
14	It is good idea to buy something through the internet auction sites or flea market apps. (オークションサイトやフリマアプリを利用して買い物するのは良いことだ)
15	Face-to-face lessons are better than online lessons. (対面授業はオンライン授業よりよい)
16	It is better to bring a lunch box than eating at university cafeterias. (大学の食堂を利用するよりもお弁当を持参する方がよい)
17	It is good for children to play video-games. (子どもがテレビゲームをするのは良いことだ)
18	Electronic books (ebooks) are better than paper books. (電子書籍は紙の本より良い)
Posttest 2	(The topics are the same as the pretest.)

The Writing Tests

To examine the longitudinal effect of timed-writing on syntactic complexity, writing fluency, and lexical diversity, timed-writing texts written at three time points were compared. The pretest was conducted at the beginning of the first semester, and approximately three months later, at the end of the first semester, posttest 1 was administered. At the end of the second semester, about five months after posttest 1, posttest 2 was administered. To counter-balance the effect of topic difficulty, the participants were randomly assigned to one of four topics. For each test, the topics were changed and they were assigned topics that differed from the previous tests. Originally, I was planning to conduct four tests, so I prepared four topics. However, I could only administer three tests to the comparison group, so the data for only three tests were used for analysis.

Pre-Task Conditions

Outlining

Planning is defined as “mental rehearsal of the message and its formulation prior to communication” (Bygate, 2001, p. 28). The participants in all three pre-task condition groups used outlining worksheets for planning (Appendix A) to generate and organize their ideas according to the format of the paragraph structure. The template for the outline, which was provided on the worksheet, included sections where they wrote their position (agree or disagree) and reasons and details to support the position. The participants could refer to the completed sheets while they engaged in the oral rehearsal pre-task and timed-writing task. Samples of the completed worksheets are shown in Appendix B.

Oral Rehearsal Pre-Task

The participants assigned to the oral rehearsal pre-task were given opportunities to rehearse their ideas orally before engaging in the timed-writing tasks. The oral rehearsal pre-task was completed in pairs after five minutes of individual planning. Referring to the outline they produced during the planning session, one student in each pair orally presented their ideas to their partner as if they were giving a speech following the structure of English writing (i.e., topic sentence, supporting ideas and details, and concluding sentence). The other person listened quietly. After two and a half minutes, the speaker and listener switched roles, and the same procedure was repeated.

Previous studies on the rehearsal of oral performances have suggested that rehearsal positively affects learners' subsequent performances of the same task (Bygate, 1996, 2001; Gass et al., 1999). Bygate (2001) explained that task repetition enhances learners' performances because the work of conceptualization, formulation, and articulation in the first trial is stored in memory and can be recycled in the second performance. In this study, the mode of the first occasion (oral rehearsal) and the second occasion (writing) differed; however, the participants could use the linguistic structures they had activated in the oral rehearsal pre-task when engaged in the timed-writing tasks. In addition, through pair work, the participants had chances to listen to the ideas of their peers and incorporate those ideas in their timed writing.

Reading Pre-Task

The participants assigned to the reading pre-task engaged with reading passages related to the topic of timed-writing task. The passages were written by me based on the style of an English textbook titled *For and Against* (Gillian & Sakamoto, 2008), which is written for Japanese university students. This textbook was chosen because each unit is focused on one problem and the viewpoints of the people who agree and disagree with the issue are represented equally. Using this format and some of the contents as a basis, I wrote the passages controlling for length and lexical level so that the participants could read and understand the texts within the given time.

Table 4 shows the Flesch-Kincaid readability score and the lexical composition. The readability score was calculated with the online calculator in Good Calculators (2023), and the frequency level of the vocabulary in the passages was calculated with

Compleat Web VP v.2.6 (Cobb, 2023). The Japanese glossary of the words whose frequency levels were at the third 1,000 high-frequency words of English or higher was attached below each passage. The passages were approximately 220 words long and a comprehension quiz, which consisted of five true-false questions, was attached to each passage. The texts and comprehension questions are shown in Appendix C.

Table 4. *Lexical Composition, Flesch-Kincaid Readability Scores, and Standard Word Units of the Reading Passages*

	Reading passage	K1 words (%)	Flesch-Kincaid score	Standard word units
1	Movie viewing at theater or at home	93.80	83.40	189.00
2	Visiting Hokkaido or Okinawa	92.30	79.60	210.00
3	Using SNS as communication tool	95.00	77.00	203.00
4	Having pets	93.20	86.30	216.00
5	Wearing school uniforms	93.30	77.70	217.00
6	Living alone or living with parents	92.70	78.00	235.00
7	Electronic or cash payments	87.20	77.00	242.00
8	Traveling by train or airplane	91.70	77.70	244.00
9	Smartphones for elementary students	85.40	70.00	206.00
10	Living in a house or an apartment	92.40	77.80	227.00
11	iPhones or Android	76.70	70.90	227.00
12	Traveling with friends or alone	94.10	76.20	227.00
13	University students doing a part-time job	93.20	74.60	265.00
14	Using internet auction sites or flea market apps.	84.60	76.20	214.00
15	Face-to-face or online lessons	88.40	70.90	266.00
16	Bring a lunch box or eating at university cafeterias	89.20	70.00	246.00
17	Children playing video-games	95.20	76.10	256.00
18	Electronic books or paper books	87.40	70.30	229.00

Note. K1 = The first 1,000 high-frequency words of English. Proper nouns were counted as K1 words; Standard word units = A word unit that consists of six characters including spaces and punctuation (Craver, 1976).

The percentage of words in the first 1,000 high-frequency band in some passages was lower than in others. For example, the percentage was 76.6% in Passage 11 (iPhone or Android smartphone). However, I found that the word *smartphone*, the equivalent of which exists in the Japanese language (*sumato fon*), is an off-list word and therefore it lowered the percentage. In addition, I conducted a pilot study with students whose English levels were the same as the participants in this study. The Japanese glossary was attached to the words they found difficult to understand. Thus, the participants did not have great difficulty understanding the passages. Table 5 shows the timing of the pre-task treatments for each task condition.

Table 5. *Timing of the Pre-Task Treatments for Each Task Condition*

Task condition	Outlining	Oral rehearsal	Reading	Timed writing	Total time
Outlining	5 minutes	--	--	12 minutes	17 minutes
Speaking	5 minutes	5 minutes	--	12 minutes	22 minutes
Reading	5 minutes	--	5 minutes	12 minutes	22 minutes

Post-Course Questionnaire

After completing Test 3, a short questionnaire asking about the timed-writing activity was conducted using Google Forms. The participants responded to three questions:

1. You have completed timed-writing with three different pre-tasks: outlining, oral rehearsal, and reading. Which task did you find most useful for completing timed-writing?
2. Please describe the reason for selecting your response in the question 1.

3. You have done timed-writing for one year. Look at your graph and briefly describe how you have changed (or not changed) over one year.

The Japanese version of the questionnaire is in Appendix D.

Procedures

The study was conducted for two academic semesters over nine months, including 18 writing sessions, one practice session, one pretest and two posttests (posttest 1, and posttest 2). The purpose of the practice sessions was to allow the participants to become used to timed-writing using their iPads. There was a six-week interval after posttest 1 because the school was closed for summer vacation. The majority of the participants had their own keyboards, but some wrote texts by directly touching the screen. Whichever way they used, I asked them to keep the same input style (i.e., with or without a keyboard) throughout the academic year when they were engaged in the timed-writing tasks. The participants completed a 12-minute timed-writing task in a practice session. After the practice session, I explained that timed-writing activities would be conducted throughout the year as a part of class activities and asked the participants if I could use their compositions as data for my research. If they agreed, they read and signed the consent form. The informed consent form is shown in Appendix E (Japanese) and Appendix F (English translation).

Table 6 shows the schedule of the study. Besides timed-writing activities, the participants in both the experimental and comparison groups read reading passages and completed reading comprehension and grammar exercises in the same textbook. They

also received instruction on general paragraph writing or essay writing. While the experimental group was engaged in timed-writing, speed reading and pronunciation exercises were used with the comparison group. Weeks 1 and 16 were course orientation days, and in Week 15 and Week 30, the participants took written final examinations. In Weeks 16 and 17, the participants engaged in activities related to sharing their experiences about summer vacation and completed textbook activities. I avoided conducting posttests on these days because the attendance rate was usually low at the beginning of the semester.

Pretest

The first timed-writing product produced after the practice session served as a pretest. All participants, including those in the comparison group, completed the timed-writing pretest task by responding to one of the following four topics: *Do you agree or disagree? A: Traveling in a foreign country better than traveling in Japan. B: Living in a city is better than living in countryside. C: Having siblings is better than being an only child. D: Japanese style breakfast is better than American style breakfast.* I randomly assigned one fourth of the participants in each class to one of the four topics. After assigning the topics, I shared document files via Google classroom (2023) that the participants could use for brainstorming, writing an outline, and the timed-writing task. First, the participants were told to spend five minutes generating and organizing their ideas. After five minutes of planning time, I announced that they should start the timed-

Table 6. *Schedule of the Study for the Experimental and Comparison Groups*

Week	Experimental group (<i>n</i> = 41)			Comparison group (<i>n</i> = 42)
	Class A (<i>n</i> = 15)	Class B (<i>n</i> = 12)	Class C (<i>n</i> = 15)	
1–2	—	—	—	—
3	Practice session	Practice session	Practice session	—
4	Pretest	Pretest	Pretest	Pretest
5–7	+ Reading	Outlining	+ Oral rehearsal	—
8–10	+ Oral rehearsal	+ Reading	Outlining	—
11–13	Outlining	+ Oral rehearsal	+ Reading	—
14	Posttest 1	Posttest 1	Posttest 1	Posttest 1
15	—	—	—	—
Summer vacation				
16	—	—	—	—
17	—	—	—	—
18–20	+ Reading	Outlining	+ Oral rehearsal	—
21–23	+ Oral rehearsal	+ Reading	Outlining	—
24–26	Outlining	+ Oral rehearsal	+ Reading	—
27	Posttest 2	Posttest 2	Posttest 2	Posttest 2
28–30	—	—	—	—

Notes. Outlining: Making an outline for 5 minutes and timed writing for 12 minutes; + Oral rehearsal = Making an outline for 5 minutes, oral rehearsal in pairs for 5 minutes, and timed writing for 12 minutes; + Reading = Reading for 5 minutes, making an outline for 5 minutes, and timed writing for 12 minutes. No activities related to timed-writing were conducted in Weeks 1–2, 15, 16–17, and 28–30.

writing task and complete it within 12 minutes. The participants had their own iPad devices. Most of the participants had their own keyboards, but some completed the timed-writing texts by touching a digital keyboard on the iPad screen. Ideally, all the participants needed to use the same input style throughout the study, so I told them to keep the same input style throughout the course when they did the timed-writing tasks. Although the type of input might have affected the length of the text they produced, as long as they used the same type of input throughout the semester, the effect on the final results was minimal. After completing the timed-writing task, the document files were collected using Google classroom.

Timed-Writing Tasks

As a part of the class activities, timed-writing sessions were implemented 18 times over nine months in the three classes assigned to the experimental group. The participants in the experimental group completed timed-writing tasks in three pre-task conditions: outlining, oral rehearsal, and reading. The pre-tasks were switched every three sessions among the three classes to counterbalance the effect of learner variables and provide equal learning opportunities. The participants assigned to the outlining condition spent five minutes completing the outlining worksheet included in the document files shared via Google classroom (2023) to brainstorm and organize their ideas. After five minutes, I told the participants to start the timed-writing task, and they spent 12 minutes to write their opinions. The participants assigned to the oral rehearsal pre-writing task also spent five minutes for planning, using the same procedures. After that, they spent another five minutes for oral rehearsal. For each session, the pairs and

which participants should speak first were randomly decided by me. Referring to their outlines produced during the planning session, one participant in each pair presented their ideas orally to their partners, and the other participants listened quietly. After two and a half minutes, they switched roles and spent another two and half minutes rehearsing (5 minutes total). Following the rehearsal session, the participants in this group engaged in the timed-writing task for 12 minutes. I told them that they were allowed to incorporate the ideas or expressions they learned from the partners during the oral rehearsal task if they wished to do so. The third pre-task, reading, required the participants to read a passage related to the day's timed-writing topic. The participants received the URL of the Google form through which they could access the passage and comprehension quizzes and were told to read and complete the comprehension quiz within five minutes. Each quiz consisted of five true-false comprehension questions that were administered to confirm that the participants had read and understood the passages. After completing the quizzes, the students submitted the forms. The answers were marked automatically and they received the results immediately after submitting the forms. After completing the reading pre-task, they spent five minutes planning. To avoid copying the texts in the passage word by word, they were told to close the Google forms that contained the reading passages before the planning session. They were told that they could use the ideas or expressions they had seen in the reading if they wished to do so. Following the planning session, the participants engaged in timed-writing for 12 minutes. After completing the timed-writing task, the participants in all groups counted the number of words using the word count function in the Google document

tool and the results were recorded in the progress check charts (see Appendix G for a sample timed writing progress chart) so that they could visually see their progress. The participants in the comparison group did not engage in timed-writing activities; they just took the pretest and the two posttests. While the participants in the experimental group were engaged in timed-writing activities, those in the comparison group were engaged in speed-reading activities. They spent approximately 20 to 25 minutes for each speed-reading activity, the equivalent length of the time spent for timed-writing in the experimental group.

Posttests

The compositions produced at the end of the first semester (posttest 1), and at the end of the second semester (posttest 2) served as the posttests in the study. The procedures for the two posttests were the same as that of the pretest. After receiving document files for the planning sheet and topics via Google classroom, they spent five minutes making an outline. Because they were using iPads, I allowed them to use an online dictionary while planning. They were then given 12 minutes to complete the timed-writing task on the document forms. After completing the timed-writing task, they counted the total number of the words and recorded the results in the progress check charts. Upon the completion of posttest 2, the participants in the experimental group responded to the post-course questionnaire. The questions are shown in Appendix D. The responses were collected online using Google forms.

Interview Sessions

Stimulated recall interview sessions were held with three volunteer participants in the experimental group to ask what they were thinking about during the pre-task sessions and how they incorporated the information on the planning sheet into their timed-writing. Only three participants volunteered, and I asked all three of them to participate in the interviews. All of the interviewees were female students. The interviews were conducted online using Zoom about one month after the final posttest. First, the participants were asked to do timed-writing again either with the oral rehearsal or reading pre-tasks. Of the three participants, the reading task was assigned to one of them and the oral rehearsal task was assigned to the other two. The participants in both pre-task groups repeated the same procedures as those used in the regular sessions. The participant assigned to the reading task received a file with a reading passage and comprehension quiz online and spent five minutes to complete the reading task, another five minutes for planning, and 12 minutes for the timed-writing task. The other two participants who engaged in the oral rehearsal task accessed Zoom simultaneously to work in pairs. After five minutes of planning time, they started the oral rehearsal pre-task, and then completed the 12-minute timed-writing task. Interview questions were asked individually to these two participants after the timed-writing session. The cameras were turned off while the participants in both pre-task groups were writing. I observed their writing via a Google document form that was shared with me during the planning and timed-writing sessions to prepare for the follow-up interview sessions. After completing the timed-writing task, I used their outlines and compositions as stimuli and

asked them to describe what they were thinking as they completed the compositions. Questions regarding each of the three pre-tasks were also asked. The interview session took about 60 minutes. The sample questions are shown in Appendix H.

Analyses

Complexity, Fluency, and Lexical Diversity Measures

All timed-writing texts written from Sessions 1 to 18 were analyzed using Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (Kyle, 2016; Lu, 2010) and Tool for the Automatic Analysis of Lexical Diversity (Kyle et al., 2020). To answer research question 1, the effect of pre-tasks on timed-writing, timed-writing samples written at four time-points were analyzed. Because a Latin-squares design was used, the participants had opportunities to engage with the three types of pre-tasks six times. Of the six sessions for each task, the texts written in the first (Time 1), third (Time 2), fourth (Time 3), and sixth (Time 4) were used as data. Table 7 indicates the measures used for the quantitative analysis.

Table 7. *Measures Used for the Complexity, Fluency, and Lexical Diversity Analyses*

	Measure
Complexity	<ul style="list-style-type: none"> • Mean length of T-unit (MLT): The ratio of the number of words to the number of T-units. • Mean length of clause (MLC): The ratio of the number of words to the number of clauses. • Dependent clauses per clause (DC/C): The ratio of the number of dependent clauses per clause.
Fluency	<ul style="list-style-type: none"> • Standard word units (SWU): Total number of standard word units (six characters including space and punctuations) produced in 12 minutes.
Lexical diversity	<ul style="list-style-type: none"> • Lexical density (LD): The ratio of the number of lexical words* to the number of words. • Moving-average type token ratio (50-word window) (MATTR): Moving-average of all segments of 50 tokens. • Measure of Textual, Lexical Diversity (MTLD): The average number of function word tokens it takes to reach a given TTR value (.720).

Lexical words*: nouns, adjectives, verbs, and adverbs (Lu, 2012).

Syntactic Complexity

I selected the indices for measuring syntactic complexity based on the complexity dimensions suggested by Norris and Ortega (2009). They proposed that syntactic complexity should be measured multidimensionally, using measures of at least three dimensions: global or general complexity, complexity by subordination, and complexity via phrasal elaboration (p. 574). According to their suggestions, I adopted mean length of T-units (MLT) for global complexity, dependent clauses per clause (DC/C) to assess complexity for subordination, and mean length of clause (MLC) for phrasal complexity. These measures are some of the most commonly used indices for

measuring syntactic complexity, and Kyle et al. (2021) indicated that mean length of T-units (MLT) and dependent clauses per clause (DC/C) are reliable indices for predicting the language development of CEFR A2 level English language learners, and one of the most commonly used indices for measuring syntactic complexity. The Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (Kyle, 2016; Lu, 2010) was used for the computations.

Writing Fluency

In this study, writing fluency was defined as “a measure of the sheer number of words or structural units a writer is able to include in their writing within a particular period of time” (Wolfe-Quintero et al., 1998, p. 14). Originally, I intended to use two measures to assess writing fluency: the total number of words produced in 12 minutes and the total number of standard word units written in 12 minutes, which is a word unit that consists of six characters including spaces and punctuation (Craver, 1976). However, the two measures were highly correlated at $r = .98$, so to avoid running multiple ANOVAs, I used the standard word unit for the fluency measure. The total number of words was dropped from the dependent variables because some researchers have argued that word length, rather than the number of total words is a more appropriate measure of the amount of language produced (Wolfe-Quintero et al., 1998). Although the use of a standard unit has been mainly discussed in reading fluency research, I used it for this study, as it is considered a more reliable measure than total number of words for writing fluency given that individual words can vary greatly in their length.

Lexical Diversity

Three indices were used to measure lexical diversity. One measure was the measure of textual, lexical diversity (MTLD; McCarthy, 2005; McCarthy & Jarvis, 2010), which showed the average number of function word tokens it takes to reach a given TTR value of .720 (Zenker & Kyle, 2021, p. 5). Another was the moving-average type token ratio (50-word window) (MATTR-50; Covington & McFall, 2010), which indicates the average type-token ratio of 50 words segments. These measures are recognized as stable indices of lexical diversity that are not affected by text length (Zenker & Kyle, 2021). The third measure was lexical density. Lu (2012), one of the developers the tool used for analysis, defined lexical density as “the ratio of the number of lexical words to the number of words” (p. 192) in which lexical words includes nouns, adjectives, verbs, and adverbs. Texts were assessed with the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (Zenker & Kyle, 2021; Lu, 2012).

Holistic Measurements

In addition to the CFL measures, two holistic measurements, content and comprehensibility, were included as dependent variables. Kuiken and Vedder (2017) emphasized the importance of considering functional adequacy (FA) when assessing L2 writing because higher CAF measures do not necessarily indicate better writing. Taking syntactic complexity as an example, non-native speakers sometimes produce long complex sentences that fundamentally ungrammatical but are pragmatically unnatural, whereas native speakers can convey the same message using shorter, syntactically

simpler sentences. Kuiken and Vedder stated that CAF measures are independent from functional adequacy and thus they should be evaluated separately. For both content and comprehensibility, I adopted the rubric created by Kuiken and Vedder because they found that the rubric was reliable even when used by non-expert raters. The descriptors for the six-point scale are shown in Table 8.

Content

The content index concerned the quantity and quality of ideas or concepts in the participants' written texts (Kuiken & Vedder, 2017). In other words, it assessed whether the essay contained adequate numbers of ideas and the degree to which the ideas were relevant to each other. The raters evaluated content using a six-point scale in which: 1 = The number of ideas is not at all adequate and insufficient and ideas are unrelated to each other, and 6 = The number of ideas is extremely adequate and they are very consistent to each other. Samples of the timed-writing texts with raters' scores are shown in Appendix I.

Comprehensibility

The comprehensibility index was focused on how much effort was required in order to understand the participants' messages (Kuiken & Vedder, 2017). This measure was used to replace the CAF accuracy measure because I found that evaluating accuracy for low-proficiency students is extremely challenging in a pilot study. This issue occurs because the number of error-free sentences is very small. More importantly, Google Classroom was the university's main learning platform, and

Table 8. Six-Point Holistic Scale

Content					
1	2	3	4	5	6
The number of ideas is <i>not at all adequate</i> and insufficient and the ideas are unrelated to each other.	The number of ideas is <i>scarcely adequate</i> and the ideas lack consistency.	The number of ideas is <i>somewhat adequate</i> , even though they are not very consistent.	The number of ideas is <i>adequate</i> and they are sufficiently consistent.	The number of ideas is <i>very adequate</i> and they are very consistent to each other.	The number of ideas is <i>extremely adequate</i> and they are very consistent to each other.
Comprehensibility					
1	2	3	4	5	6
The text is <i>not at all comprehensible</i> . Ideas and purposes are unclearly stated and the efforts of the reader to understand the text are ineffective.	The text is <i>scarcely comprehensible</i> . Its purposes are not clearly stated and the reader struggles to understand the ideas of the writer. The reader has to guess most of the ideas and purposes.	The text is <i>somewhat comprehensible</i> . Some sentences are hard to understand at a first reading. A second reading helps to clarify the purposes of the text and the ideas conveyed, but some doubts persist.	The text is <i>comprehensible</i> . Only a few sentences are unclear but are understood, without too much effort, after a second reading.	The text is <i>easily comprehensible</i> and reads smoothly. Comprehensibility is not an issue.	The text is <i>very easily comprehensible</i> and highly readable. The ideas and the purposes are clearly stated.

From "Functional adequacy in L2 writing: Towards a new rating scale," by F. Kuiken and I. Vedder, 2017, *Language Testing*, 34(3), p. 335 (<https://doi.org/10.1177/0265532216663991>). Adapted with permission.

the participants used Google Docs for their writing, which is equipped with an automatic error correction function. Thus, counting the number of errors might not have provided valid data. Instead of focusing on the number of local errors, measures that could evaluate accuracy more globally were more appropriate for this study.

Content and comprehensibility were evaluated by five raters, including myself, using the six-point scale rubric created by Kuiken and Vedder (2017) (Table 8). Google forms was used to collect responses. The raters were teachers with more than ten years of teaching experience at the university level who held Masters or Ph.D. degrees in TESOL or applied linguistics. Before they began scoring, they participated in a Zoom training session individually. I first explained the evaluation criteria for content and comprehensibility and showed sample writings for each score. The raters were then given about 10 to 15 minutes to randomly evaluate six sample writings that varied in quality. The results were then compared with my ratings. We had discussions about any inconsistent ratings until we reached an agreement. Each training session took approximately 40 to 60 minutes.

The compositions written in Weeks 7, 10, and 13 (Topics 3, 6, and 9) were used as the data. The pre-tasks were rotated among the three classes every three weeks. To reduce the burden on the raters, each rater evaluated about 60% of the data. Rater bias was controlled for using FACETS (Linacre, 2023), and the Rasch Fair Average was used for the analyses. According to Linacre (2023), the Fair Average is the raw score adjusted for its context. Its calculation is as follows.

“Fair Average” for each element of the target facet is computed and calculated

as though every observation of that element in the data was for an average element of every other facet. Then these artificial observations are averaged by the number of observations of the target element. (p. 309)

For example, if participant A receives a rating of 3 from a lenient rater and participant B receives a rating of 3 from a severe rater, interpreting their scores as the same value is unfair due to the difference in the raters' severity. In the Fair Average, participant A's score is transformed to 2.5 and participant B's score is 3.5; thus, this gives a fair comparison. By using the Fair Average, it is possible to compare the scores evaluated by raters of different severity.

Interviews and Post-Course Questionnaire

The interview sessions were conducted using Zoom. The sessions were recorded using its recording function and then the interviews were transcribed for analysis. The purpose of the interviews was to obtain data that complemented the results of the quantitative analyses. I reviewed the transcripts and made annotations on the comments or excerpts that could explain or provide additional explanations of the quantitative results. I particularly focused on the comments explaining the participants' mental activities during the pre-tasks and how they took advantages of each pre-task activity. The data were used for triangulation together with other qualitative data such as the post-course questionnaire and the word similarity rate between the reading passage and timed-writing text.

Additional Text Analysis on Lexical Items

To investigate the effect of the reading pre-tasks on lexical diversity, an additional text analysis was conducted. Specifically, the lexical items used in the reading passages and the compositions written by the reading pre-task group were compared using Text Lex Compare (Cobb, 2023), an online tool that computes the similarity of lexical items between two texts. The text of the sample passage and the compositions written in Weeks 7, 10, and 13 were used as the data, and a similarity rate was calculated separately for each session. The similarity rate of the other two groups that did not read the passage—the outlining and oral rehearsal groups—were also included in the analysis as baseline data.

CHAPTER 4

PRELIMINARY ANALYSES

The Overview of the FACETS Analysis

In order to obtain fair average scores that were used for the inferential statistics, the scores of the content and comprehensibility evaluated by the raters were analyzed using FACETS (Linacre, 2023), a MFRM (Many-Facet Rasch Measurement) computer software program. Before reporting the results of the inferential statistics, findings from the FACETS analysis are presented. There were five facets in the analysis: the participants (44 elements), raters (5 elements), topics (3 elements), task type (3 elements), and rating criteria. The criteria consisted of six categories: Content outlining, Comprehensibility outlining, Content oral rehearsal, Comprehensibility oral rehearsal, Content reading, and Comprehensibility reading. Figure 3 shows each facet and how they are related to each other. Starting from the leftmost column, it shows the logit measures, the raters, the participants, the task types, and the evaluation criteria. I focus on the fifth column, evaluation criteria, which is directly relevant to this study. Table 9 shows the Rasch statistics for the six criteria. The standardized infit statistics of five out of six criteria fit the acceptable range of -2.00 to +2.00, but the comprehensibility score of the oral rehearsal pre-task was 2.40, which indicated underfit to the Rasch model. The item difficulty for each criterion ranged from -0.24 to 0.16 logits. For both content and comprehensibility, the outlining pre-task condition was the most difficult, the oral rehearsal condition was second, and the reading condition was the easiest. The difference between oral rehearsal and reading for content was small.

Figure 3. *Wright Map for Content and Comprehensibility*

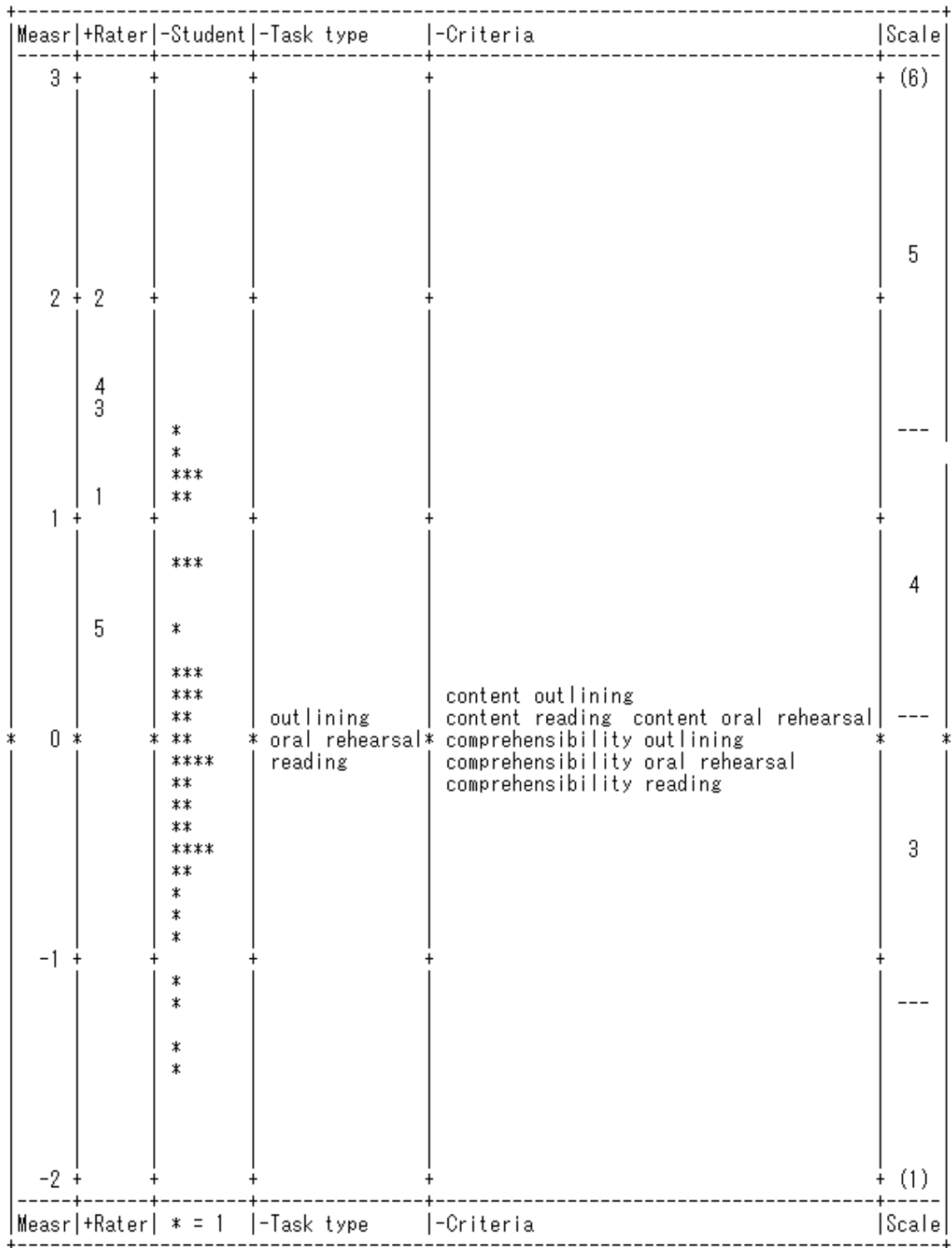


Table 9. *Rasch Statistics for the Six Evaluation Criteria*

	Rasch measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Output ZSTD	PMCorr
Comp. R	-0.24	0.10	0.89	-1.00	0.87	-1.10	.67
Comp. OR	-0.13	0.10	1.31	2.40	1.29	2.30	.42
Comp. P	0.02	0.09	0.97	-0.20	0.95	-0.40	.66
Cont. R	0.09	0.10	0.94	-0.50	0.93	-0.60	.57
Cont. OR	0.10	0.10	1.05	0.40	1.05	0.40	.56
Cont. P	0.16	0.09	0.92	-0.70	0.93	-0.60	.65

Note. Comp = Comprehensibility; Cont = Content; R = Reading; OR = Oral Rehearsal; O = Outlining; PMCorr = Part-measure correlation

Table 10 shows the category statistics. The fifth column from the left indicates the average measure, which can be used to confirm whether the rating scale functioned well. For well-functioning rating scale, the average measures increase as the variable increases, which indicates those with higher ability obtain higher scores (Bond & Fox, 2007). As Table 10 indicates, this pattern was not violated; thus, the average measures of the scale functioned appropriately.

Another way to assess the rating scale is to inspect the thresholds or step calibrations, which are shown in the column located second from the right. Similar to the average measures, it should increase in a stepwise fashion and the difference between adjacent thresholds should be no less than 1.40 logits and no more than 5.00 logits to avoid large gaps in the variable (Linacre, 1999). Table 10 shows that the thresholds increase stepwise, but the distance between some adjacent categories is below 1.40 (Category 3–4 = 0.58; Category 4–5 = 1.36; Category 5 and 6 = 1.15), which indicates that the distance between these adjacent categories are so close that the steps are not entirely clear.

Table 10. *Category Statistics for the Six-Point Rating Scale*

Category	Count (%)	Cum. %	Average measure	Expected measure	Outfit MNSQ	Measure	SE
1	2 (0%)	0%	-0.06	-0.35	1.20	—	—
2	45 (5%)	5%	0.14	0.03	1.20	-3.29	0.71
3	121 (13%)	19%	0.51	0.55	0.90	-0.70	0.16
4	315 (35%)	53%	1.09	1.12	0.90	-0.12	0.10
5	287 (32%)	85%	1.68	1.64	0.90	1.48	0.08
6	134 (15%)	100%	2.06	2.07	1.00	2.63	0.10

Note. Cum. % = Cumulative percent.

The last criterion concerns the rating scale fit statistics. According to Linacre (2023), outfit MNSQ values between 0.50 and 1.50 are productive for measurement. All the outfit MNSQ statistics for the six categories were within this range.

Comprehension Quizzes for Reading Passages

The participants in the reading group took five reading comprehension questions after reading. These quizzes were conducted to make sure that they understood the content of the passages adequately and to encourage them to engage in the reading task seriously. Table 11 shows the percentage of the correct answers for each quiz. The results varied from 75% to 100%, which suggests that some passages were more difficult to understand than others. However, the results also showed that for some passages with lower scores (Passages 6, 7, and 13), particular participants who got low scores lowered the average. Of the five questions for each quiz, the participants answered four or five questions correctly on average.

Table 11. *The Percentage of Correct Answers for Each Comprehension Quiz*

Passage	1	2	3	4	5	6	7	8	9
%	85.3	100.0	97.3	96.0	96.7	78.7	75.7	85.3	86.7
Passage	10	11	12	13	14	15	16	17	18
%	98.7	94.7	96.3	77.3	98.0	89.1	97.3	82.5	100.0

CHAPTER 5

RESULTS

In this Chapter, I first introduce the results for the first research question concerning the effectiveness of three types of pre-writing tasks—outlining, oral rehearsal, and reading—on writing quality, which was operationalized as syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility. After that, I show the results regarding the second research question regarding how the participants made use of each pre-writing task and how it was reflected in the quality of their written products. Specifically, the findings from the interviews and additional text analyses are reported. Finally, I introduce the findings for the longitudinal investigation of how fluency-first, meaning-focused writing activities contributed to writing development.

The Effect of the Three Pre-Tasks on Timed-Writing

Table 12 shows the descriptive statistics of three syntactic complexity measures—mean length T-unit (MLT), mean length clause (MLC), and dependent clause per clause (DC/C)—for each pre-task at four time points. The growth patterns of MLT and MLC are similar. The scores for MLT and MLC increased from Time 2 to 3 and were the highest at Time 3. On the other hand, the score for DC/C developed from Time 3 to 4, and was the highest at Time 4, showing the opposite pattern of MLT and MLC.

Table 12. *Descriptive Statistics for the Syntactic Complexity Measures for Each Pre-Task*

	Mean length T-unit (MLT)											
	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	10.53	10.31	10.94	10.15	9.90	10.08	10.92	10.68	9.60	10.19	11.16	10.44
<i>SD</i>	2.41	2.13	2.18	2.16	1.89	1.56	1.79	1.50	2.07	1.67	2.12	1.49
Skewness	1.06	1.55	1.76	1.69	0.22	0.74	1.75	1.11	1.86	0.78	0.52	0.50
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	0.74	3.70	3.47	6.61	-0.72	1.60	6.22	2.00	5.07	0.78	-0.73	0.87
SEK	0.72	0.72	0.72	0.72	0.73	0.73	0.73	0.73	0.72	0.72	0.72	0.72

	Dependent clause per clause (DC/C)											
	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	0.31	0.29	0.30	0.31	0.29	0.30	0.30	0.32	0.29	0.29	0.32	0.33
<i>SD</i>	0.10	0.10	0.09	0.12	0.11	0.10	0.10	0.07	0.09	0.11	0.09	0.10
Skewness	-0.22	0.19	0.47	0.16	0.05	-0.23	0.33	0.28	-0.04	0.08	-0.63	0.14
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	-0.15	-0.11	1.31	0.09	0.74	1.48	-0.07	2.62	-0.10	0.92	0.00	-1.10
SEK	0.72	0.72	0.72	0.72	0.73	0.73	0.73	0.73	0.72	0.72	0.72	0.72

	Mean length clause (MLC)											
	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	6.97	7.19	7.64	6.76	6.89	7.12	7.53	7.31	6.65	7.28	7.40	6.82
<i>SD</i>	0.89	1.00	0.99	0.63	1.00	1.04	0.98	0.85	0.92	0.93	0.97	0.74
Skewness	-0.05	1.01	2.61	0.14	0.56	0.74	0.41	2.07	0.87	0.01	0.48	-0.32
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	-1.00	1.64	9.84	1.18	0.14	0.42	0.33	5.48	1.47	-0.96	0.10	-0.05
SEK	0.72	0.72	0.72	0.72	0.73	0.73	0.73	0.73	0.72	0.72	0.72	0.72

Note. Outlining *n*-size = 41; Oral rehearsal *n*-size = 40; Reading *n*-size = 41.

Figure 4. *Changes in Mean Length of T-Unit*

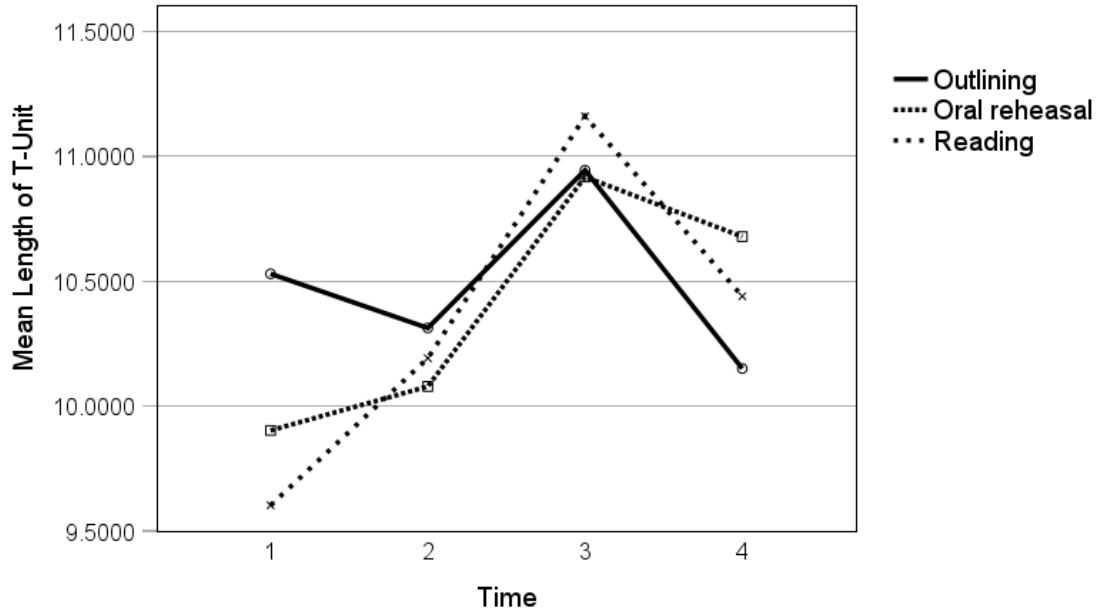


Figure 5. *Changes in Dependent Clauses per Clause*

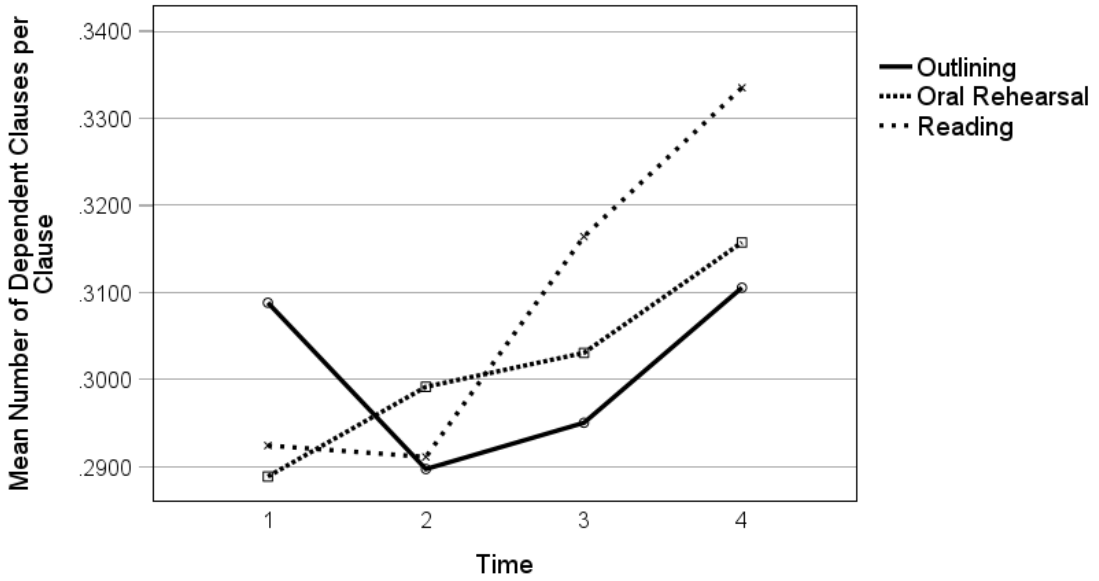
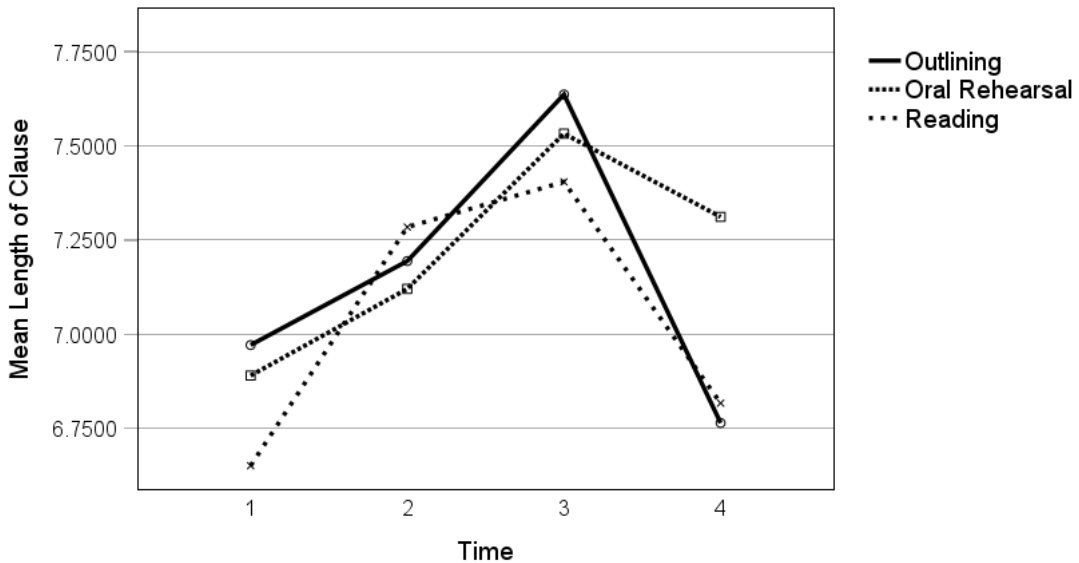


Figure 6. *Changes in Mean Length of Clause*



To answer the first research question concerning the effect of the three planning conditions—outlining, oral rehearsal, and reading—on the three syntactic complexity measures, mean length T-unit (MLT), dependent clause per clause (DC/C), and mean length of clause (MLC), a mixed MANOVA was conducted. The dependent variables were time with four levels (Times 1, 2, 3, and 4) and syntactic complexity measures with three levels (MLT, DC/C, and MLC). The independent variable was task type with three levels—outlining, oral rehearsal, and reading. Because the ANOVA was run three times for each dependent variable, a Bonferroni procedure was used to control Type I error. The alpha was set to .025 ($.05/3 = .025$).

After removing four outliers from the dataset, the normality assumption was checked before running the test. The distributions of z -kurtosis for some variables exceeded the $< |2.58|$ criterion and therefore violated the normal distribution

assumption. Thus, the results need to be interpreted with caution. High kurtosis can lead to low statistical power.

Before running a mixed MANOVA, the sphericity assumption was checked. The Mauchly's test of sphericity showed that the assumptions for MLT and DC/C were met (MLT: Chi-square = 3.79, $p = .58$; DC/C: Chi-square = 7.03, $p = .22$). However, MLC violated the sphericity assumption, Chi-square = 17.67; $p = .003$, so the Greenhouse-Geisser results are reported for MLC. There was a significant main effect of time on MLT, $F(3, 357) = 11.29, p < .001$, partial $\eta^2 = .09$, and MLC, $F(2.88, 329.34) = 15.09, p < .001$, partial $\eta^2 = .11$, but the time effect for DC/C was insignificant after applying the Bonferroni correction, $F(3, 357) = 2.57, p = .05$, partial $\eta^2 = .02$. Contrasts revealed that there was a significant difference between Time 1 and Time 3 for MLT, $F(1, 119) = .89, p < .001$, partial $\eta^2 = .19$, and a significant difference between Time 1 and Time 2, $F(1, 119) = 9.51, p = .003$, partial $\eta^2 = .07$, and between Time 1 and Time 3, $F(1, 119) = 34.50, p < .001$, partial $\eta^2 = .23$ for MLC. These results indicated that the scores for both indices developed between Time 1 and Time 3, but the growth declined between Time 3 and 4. There was no main effect of task type on any of the three syntactic complexity measures [MLT: $F(2, 119) = .09, p = .92$, partial $\eta^2 = .001$; DC/C: $F(2, 119) = .14, p = .87$, partial $\eta^2 = .002$; MLC: $F(2, 119) = 1.07, p = .35$, partial $\eta^2 = .02$]. Moreover, there was no interaction between time and task type on the three syntactic complexity measures [MLT: $F(6, 357) = 2.06, p = .06$, partial $\eta^2 = .03$; DC/C: $F(6, 357) = .69, p = .66$, partial $\eta^2 = .01$; MLC: $F(5.54, 329.35) = 1.99, p = .07$, partial $\eta^2 = .03$]. These results indicated that there were changes in the syntactic

complexity scores as the time progressed, but task type did not significantly impact the syntactic complexity of the timed-writing products.

Table 13 shows the descriptive statistics for the standard word unit for each pre-task at the four time points. The scores increased steadily for all three pre-task groups between Time 1 and Time 2, and Time 2 and Time 3, but they leveled off between Time 3 and Time 4. At Time 1, the figure for the oral rehearsal group was the lowest, but it developed sharply as the time progressed. At Time 4, the oral rehearsal group outperformed the other two groups. The figures for the reading pre-task were not as high as those of the oral rehearsal or outlining groups throughout the intervention period.

To examine the effect of the three planning conditions—outlining, oral rehearsal, and reading—on fluency, standard word units, a mixed ANOVA was conducted. The dependent variable was the values of the standard word units for Times 1, 2, 3, and 4. The independent variable was task type with three levels—outlining, oral rehearsal, and reading. After removing five outliers from the dataset, the normality assumption was assessed by checking z -skewness and z -kurtosis before running the test. The normality assumption was met.

Before running the mixed ANOVA, the sphericity assumption was checked and met, $\text{Chi-square} = 8.32, p = .14$. There was a significant main effect of time on SWU, $F(3, 354) = 46.44, p < .001$, partial $\eta^2 = .28$. Contrasts revealed that there was a significant difference between Time 1 and Time 2, $F(1, 118) = 36.93, p < .001$, partial $\eta^2 = .24$, Time 1 and Time 3, $F(1, 118) = 101.34, p < .001$, partial $\eta^2 = .46$, and Time 1

and Time 4, $F(1, 118) = 90.26, p < .001, \text{partial } \eta^2 = .43$. However, there was no significant effect of task type $F(2, 118) = .41, p < .66, \text{partial } \eta^2 = .007$ or interaction effect $F(6, 354) = 1.34, p = .24, \text{partial } \eta^2 = .02$ between SWU and task type. This result indicated that fluency developed over time but task type did not significantly impact the timed-writing fluency measure.

Figure 7. Changes in the Number of Standard Word Units per 12 Minutes

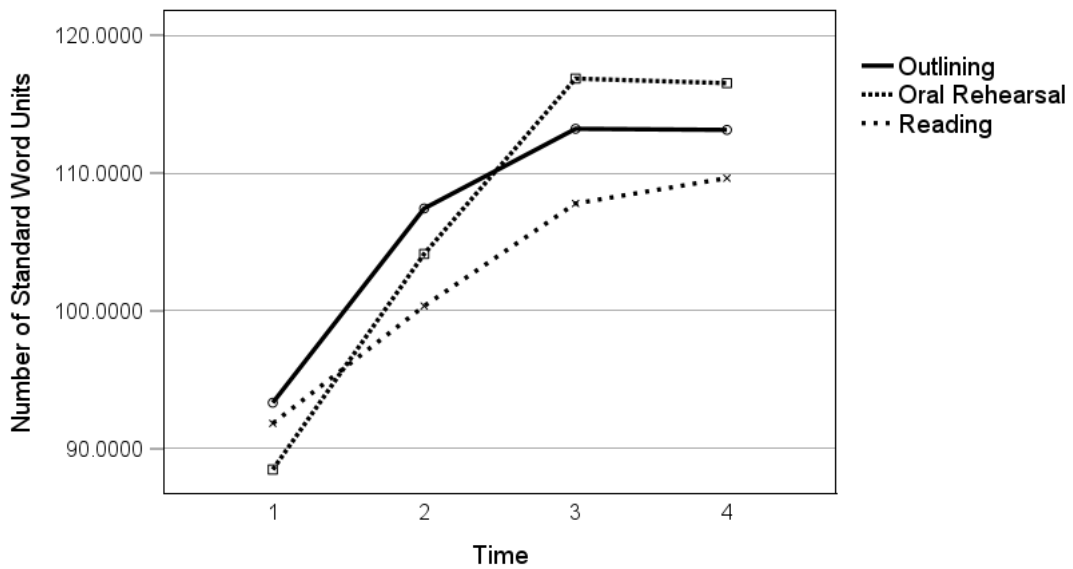


Table 13. *Descriptive Statistics for Standard Word Units for Each Pre-Task*

	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	93.34	107.43	113.21	113.13	88.49	104.14	116.85	116.52	91.84	100.35	107.79	109.63
<i>SD</i>	29.15	31.19	28.91	28.34	20.19	31.13	27.51	28.30	27.62	26.20	27.47	28.25
Skewness	1.30	0.38	0.31	0.16	0.80	0.84	0.83	0.45	0.28	-0.01	-0.07	0.19
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.38	0.38	0.38	0.38
Kurtosis	2.35	-0.17	-0.06	0.80	0.45	1.15	1.06	0.07	-0.20	-0.65	-0.64	0.30
SEK	0.72	0.72	0.72	0.72	0.73	0.73	0.73	0.73	0.74	0.74	0.74	0.74

Note. Outlining *n*-size = 42; Oral rehearsal *n*-size = 42; Reading *n*-size = 42.

Table 14 shows the descriptive statistics for the three lexical diversity measures for each pre-task at the four time points. The trajectories of MATTR, and MTLTLD are similar. The peak was at Time 2, and the figures decreased after that time point. In the case of MATTR and MTLTLD, the reading group outperformed the other two groups. On the other hand, lexical density reached the highest point at Time 3 and the oral rehearsal group outperformed the other two groups.

To explore the effect of the three planning conditions—outlining, oral rehearsal, and reading—on three lexical diversity measures, lexical density (LD), moving-average type token ratio (50-word window) (MATTR), and measure of textual, lexical diversity (MTLTD), a mixed MANOVA was conducted. The dependent variables were time with four levels (Times 1, 2, 3, and 4) and lexical diversity measures with three levels (LD, MATTR, and MTLTD). The independent variable is task type with three levels—outlining, oral rehearsal, and reading. Because the ANOVA was run three times for each dependent variable, a Bonferroni procedure was used to control for Type I error. The alpha was set to .025 ($.05/3 = .025$). There were no outliers, and the z -skewness and z -kurtosis statistics showed that the data were normality distributed.

Before running the mixed-MANOVA, the sphericity assumption was checked. Mauchly's test of sphericity showed that the figures for all three measures were insignificant (LD: Chi-square = 6.49, $p = .26$; MATTR: Chi-square = 5.60, $p = .35$; MTLTD: Chi-square = 10.21, $p = .07$). There was a significant main effect of time on all three measures (LD: $F[3, 369] = 7.80$, $p > .001$, partial $\eta^2 = .06$; MATTR: $F[3, 369] = 23.57$, $p > .001$, partial $\eta^2 = .16$, and MTLTD: $F[3, 369] = 27.56$, $p > .001$, partial η^2

= .18). Contrasts revealed that there was a significant difference between Time 1 and Time 2 for two out of the three measures (MATTR: $F[1, 123] = 41.47, p < .001$, partial $\eta^2 = .25$; MTLT: $F[1, 123] = 49.61, p < .001$, partial $\eta^2 = .29$), and between Time 1 and Time 3 for all three measures (LD: $F[1, 123] = 20.04, p < .001$, partial $\eta^2 = .14$; MATTR: $F[1, 123] = 24.25, p < .001$, partial $\eta^2 = .17$; MTLT: $F[1, 123] = 24.38, p < .001$, partial $\eta^2 = .17$). There was no significant difference between Time 1 and Time 4 for any of the three lexical diversity measures, indicating that the scores for three lexical diversity measures developed between Time 1 and Time 3, but the growth declined between Times 3 and 4, showing a similar pattern as syntactic complexity. There was no significant effect of task type on the three lexical diversity measures and no interaction between time and task type. These results indicated that task type or the combination of task type and time did not significantly impact the three lexical diversity measures of the timed-writing task.

Table 14. *Descriptive Statistics of the Lexical Diversity Measures for Each Pre-Task*

	Lexical density (LD)											
	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	0.52	0.52	0.54	0.53	0.52	0.55	0.56	0.53	0.52	0.53	0.54	0.53
<i>SD</i>	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.06	0.05	0.05	0.04	0.04
Skewness	-0.58	-0.26	0.40	-0.41	-0.08	-0.70	-0.08	-0.07	-0.68	0.39	-0.48	0.60
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	0.13	-0.60	0.82	0.72	0.99	1.09	0.22	-0.62	0.39	0.72	0.33	-0.11
SEK	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72

	Moving-average type-token ratio (MATTR)											
	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	0.67	0.71	0.70	0.67	0.68	0.71	0.71	0.68	0.69	0.72	0.71	0.69
<i>SD</i>	0.06	0.05	0.05	0.05	0.05	0.04	0.05	0.04	0.06	0.05	0.04	0.05
Skewness	-0.13	0.10	-0.33	0.08	-0.27	0.26	-0.59	-0.35	-0.48	0.20	-0.45	-0.01
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	0.03	-0.54	1.17	1.27	-0.15	0.05	0.44	0.57	0.12	-0.48	1.27	-0.91
SEK	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72

	The measure of textual lexical diversity (MTLD)											
	Outlining				Oral rehearsal				Reading			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
<i>M</i>	41.02	51.11	46.06	39.26	41.94	50.01	47.91	41.16	41.98	51.86	49.01	43.40
<i>SD</i>	14.13	13.65	9.64	10.01	11.73	8.70	12.54	7.50	9.65	15.04	10.13	12.43
Skewness	2.19	1.01	0.86	0.66	1.25	0.77	0.89	1.01	0.25	1.17	0.32	1.07
SES	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	6.47	0.53	3.33	1.91	2.26	0.39	0.47	3.07	-1.00	2.09	-0.21	1.86
SEK	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72

Note. Outlining *n*-size = 42, Oral rehearsal *n*-size = 42, Reading *n*-size = 42.

Figure 8. Changes in Lexical Density

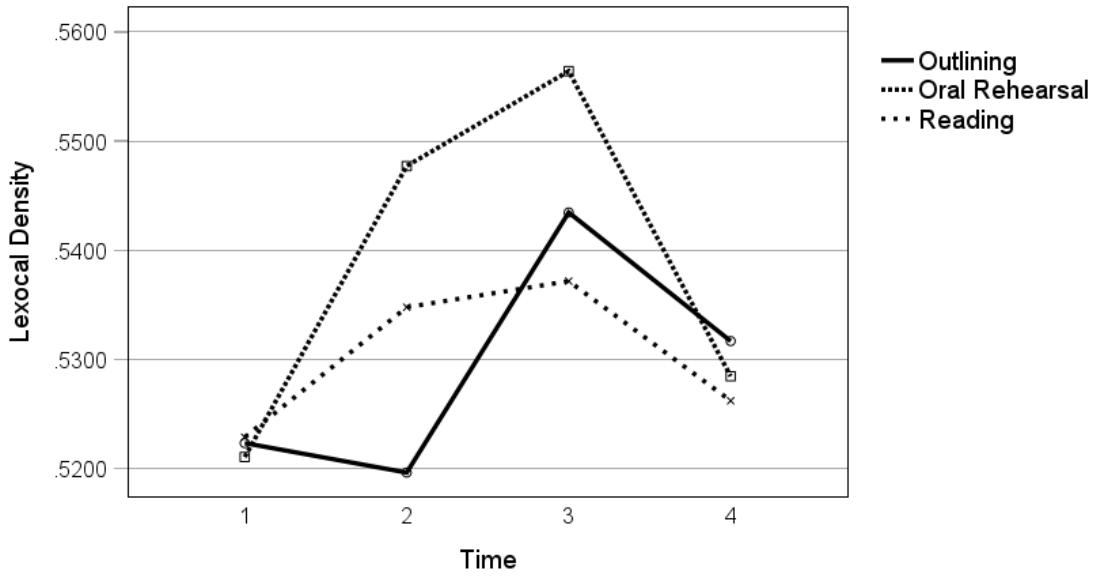


Figure 9. Changes in the Moving-Average Type Token Ratio (MATTR)

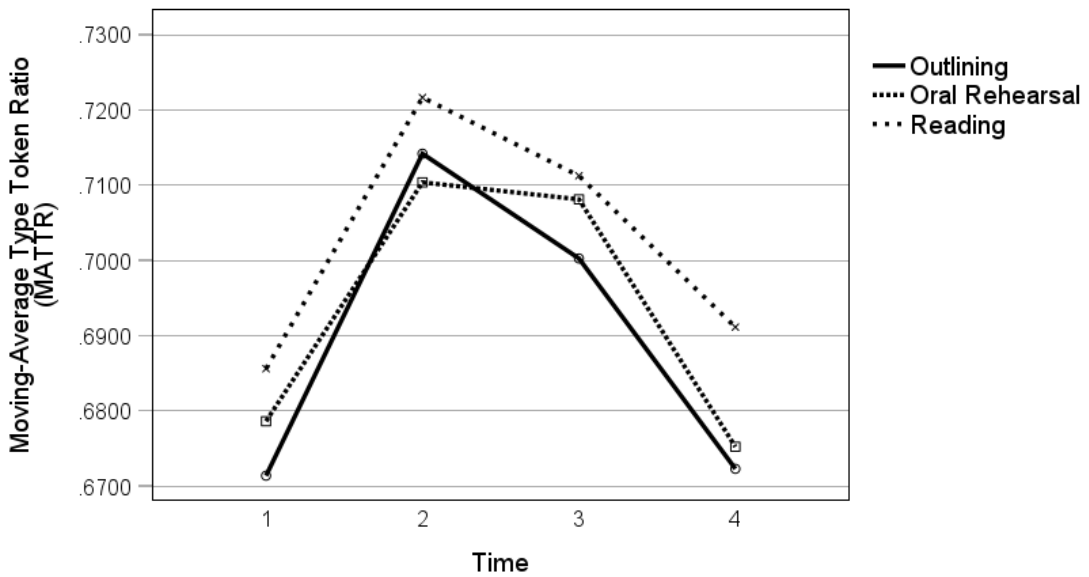


Figure 10. *Changes in the Measure of Textual Lexical Diversity (MTLD)*

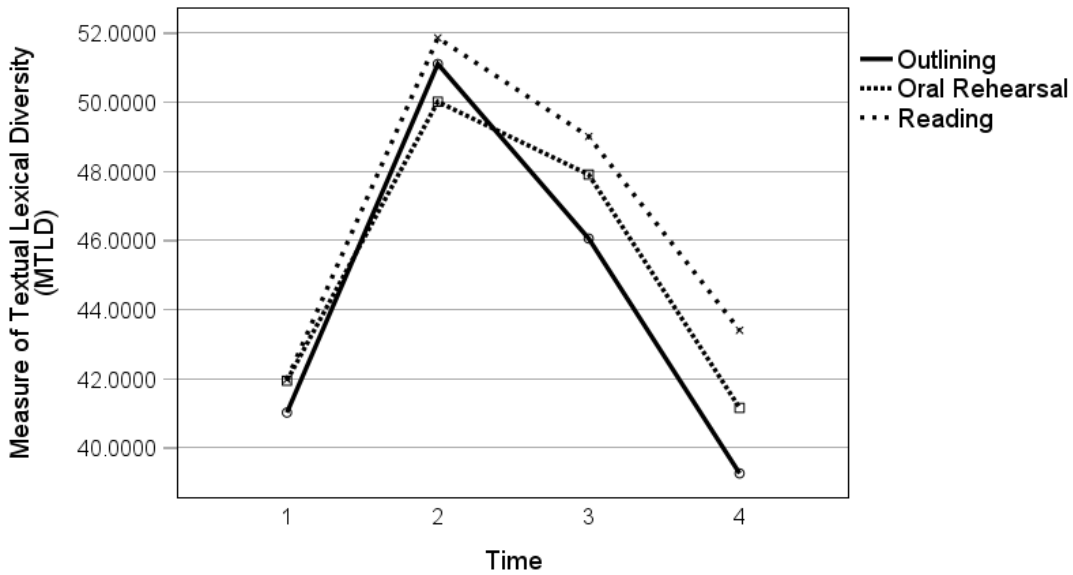


Table 15 shows the descriptive statistics for content and comprehensibility for each pre-task at three time points. The content scores of the oral rehearsal re-task was slightly higher than those for the reading pre-task, but the scores were almost the same, and they outperformed outlining pre-task. For comprehensibility, the score of the reading pre-task was the highest, and that of the outlining pre-task were the lowest.

For the two holistic measures, content and comprehensibility, one-way ANOVAs are conducted separately for each dependent variable. The dependent variables were content and comprehensibility scores for week 7, 10, and 13 (topic 3, 6, and 9) and independent variables were task type with three levels (outlining, oral rehearsal, and reading). The Levene's test for both content and comprehensibility scores were insignificant (Content: $F[2, 123] = .27, p = .77$; Comprehensibility: $F[2, 123] = 1.08, p = .34$). Two one-way ANOVAs were run for each of the content and comprehensibility scores and the test results were both insignificant (Content: $F[2, 123] = .65, p = .53$,

partial $\eta^2 = .01$; Comprehensibility: $F[2, 123] = 2.65, p = .07, \eta^2 = .04$). The type of pre-task did not have significant effect on content or comprehensibility.

Table 15. *Descriptive Statistics for the Content and Comprehensibility Measures for Each Pre-Task*

	Content			Comprehensibility		
	Oral			Oral		
	Outlining	Rehearsal	Reading	Outlining	Rehearsal	Reading
<i>M</i>	4.18	4.37	4.36	4.42	4.67	4.74
<i>SD</i>	0.91	0.82	0.87	0.80	0.63	0.62
Skewness	0.01	-0.57	-0.46	-0.71	0.00	-0.23
SES	0.37	0.37	0.36	0.37	0.37	0.36
Kurtosis	-0.71	-0.09	-0.30	0.55	-0.60	-0.35
SEK	0.72	0.73	0.70	0.72	0.73	0.70

Notes. Content and Comprehensibility scores are the Rasch fair average scores of the six-point scales evaluated by raters. Outlining n -size = 42; Oral rehearsal n -size = 40; Reading n -size = 44.

To summarize the main points, overall, the type of pre-task did not have a significant effect on the syntactic complexity, lexical diversity, writing fluency, content, and comprehensibility of the immediate timed-writing tasks. In contrast, the effect of time on syntactic complexity, writing fluency, and lexical diversity measures was significant, indicating an increase in the scores between Time 1 and Time 3. However, in the case of the syntactic complexity and lexical diversity measures, the scores decreased between Time 3 and Time 4, and the difference between Time 1 and Time 4 was not significant. There was no significant effect of time on content and comprehensibility.

The Roles of Pre-Tasks

To answer the second research question concerning how the participants made use of each pre-task condition, how it was reflected in the quality of their written products, and to triangulate the findings from the quantitative analyses, post-course questionnaires and the stimulated recall interviews were analyzed. In addition, to examine the effect of the reading passage on the lexical diversity of the timed-writing products, the word similarity rate between the reading passage and the timed-writing texts written by the reading pre-task group was calculated.

Post-Course Questionnaire

To supplement the qualitative data, I report the results of the post-course questionnaire administered to the participants in the experimental groups after Test 4. The questions were: (a) Which of the three pre-tasks—outlining, oral rehearsal, and reading—did you find most useful for timed-writing? (b) Please describe the reason for selecting your response in the question 1. Table 16 shows the results of these queries.

Nineteen participants (50%) selected the reading pre-task as most helpful for improving writing fluency. Sixteen comments indicated that the reading pre-task improved their writing fluency because it helped them generate ideas, and eight participants indicated that they learned words and multi-word expressions from the reading passage. The oral rehearsal pre-task, which was selected by 15 participants (40%), helped them generate ideas and some participants commented on the positive effect of the task for organizing ideas and for having a rehearsal effect. The outlining task was selected by four participants (10%), one of whom responded that this pre-task was the most helpful for generating ideas. The results of the interviews and questionnaire

contradicted one another because the questionnaire responses indicated that both the oral rehearsal and reading pre-tasks helped the participants generate ideas while the interview participants gave opposite answers. This point will be discussed in the chapter 6.

Table 16. *Most Helpful Timed-Writing Pre-Tasks and the Reasons*

Pre-task	Categories	Number of comments	Example comments
Outlining <i>n</i> = 4 (10.5%)	Generate new ideas	1	I thought just simply making an outline helped me most to generate ideas.
	Others	3	I had time to think before writing.
Oral rehearsal <i>n</i> = 15 (39.5%)	Generate new ideas	6	I could use my friend's ideas for my writing. As I made effort to continue speaking, sometimes new ideas that were not in the outline came to my mind.
	Organize the ideas	6	I could organize my ideas as recite them aloud.
	Rehearsal	3	It was easy to write because my ideas are already verbalized in the oral rehearsal task.
	Verbalize ideas	3	Speaking helped me to verbalize my ideas in English.
	Check my English	2	I could check my English as I said my ideas aloud.
Reading <i>n</i> = 19 (50.0%)	Others	1	I was able to expand my perspective as I listened to my friends.
	Generate new ideas	16	I could know other people's opinions and it helped me to generate my own ideas. Reading passage presented the opinions of the people from both sides, agree and disagree. So, I could develop my ideas from there and it was easier to write the outline.
	Words & grammar	8	I could learn new words and expressions in the reading and use it for my writing. I could see English grammar structures in advance. So, it was easier for me to express my ideas with my own words.
	Paragraph structure	2	I could learn how to develop my idea.
	Others	5	It helped me to understand the intention of the task.

Note. *The number Includes comments from the same students.

Interviews

Stimulated recall interviews were conducted with three female volunteers. The participants were recruited via email, and I asked all of three of them who contacted me to participate in the study. The four questions asked in the interviews were as follows: (a) Using the outlining worksheet and composition written just before the interview, please explain what you did and what you were thinking step by step; (b) Did you use any ideas or words in the reading text when you made an outline? Which ones?; (c) Did you use any ideas or expressions that you heard from your partner in your writing? Which ones? (d) In class, you had opportunities to engage in timed-writing with three pre-task conditions: outlining, oral rehearsal and reading. Which one did you find most useful for timed-writing? Why do you think so? The questions (a), (b), and (c) were concerned with the timed-writing activity conducted just before the interviews. The participants were asked to recall their experiences of writing an opinion paragraph for the topic *Do you agree or disagree? It is good for elementary school students to have smartphone?* The last question was a reflection on the timed-writing activities completed throughout the academic year. The interviews were conducted in Japanese.

The first question concerned what they did during the 5-minute planning time period. The Google document files used for making the outlines were shared with me via a cloud service. This allowed me to observe the process of the participants completing the outline in real-time. I observed the process focusing on how they generated ideas, wrote a topic sentence, and listed three supporting points. All three of them wrote a topic sentence first, and then worked on three supporting points. They struggled to express their ideas in the supporting points in English, and after writing the three supporting

ideas, they ran out of time. Their outlines also showed that they had little time to write supporting details.

[Selecting three reasons] P1 = Participant 1; P2 = Participant 2; P3 = Participant 3; I = The interviewer (myself)

1. P1: During the five minutes, there wasn't enough time, so I omitted subjects. (...) If I had only two supporting ideas, I would have to write a lot [of supporting details for each point], so I first tried to think of three ideas. To begin with, I wrote I agree with the topic, and for the first reason, I wrote elementary school students can find out what they don't know in their daily lives easily if they have smart phones.

5分間のところは、時間ないんで、まずわかる主語は省略しようって思っ
て。(...)理由が2つとか1つだとすごい分量をいっぱい考えなきゃいけなく
なるから、とりあえず3つあげようと思って。まず私はトピックに賛成しま
すって言うのを書いて、賛成する理由としてまず1つ目が、小学生がスマホ
持ったら生活の中でわからないことがあったときに簡単に調べられるって書
いたんですけど。

2. P2: I had to decide my position first, and I wrote "I agree." Then, two reasons came up, "[parents can] check the safety of their children" and "people should learn how to use the Internet while they are children because it is going to be a necessary skill". (...) I also wanted to say "[children can] learn many things efficiently, but I could not find words to express that idea.

I: So, you had another idea but ran out of time before you translated it in English.

P2: Yes.

P2: まずは賛成か反対かを考えて、ぱっと「賛成」が出てきたので、「I agree」
で文を書いてから、2つの理由が出てきました。安全を確認できるって言う
のと、インターネットが必ず必要になるからやるなら小さいうちにやってお
いた方がいいって言うことを書きたくて。(...)あと、効率よくいろんなこと
を学習できるっていうことを言いたかったんですけど、そこら辺の単語が全
然出てこなくて。

I: じゃあ、もう1つ考えがあったけど、それを英語にする前に時間が終わっ
てしまったんですかね。

P2: はい。

3. P3: In five minutes, I thought I would write three reasons first and then add details later. (...) For the third reason, though there was not enough time to complete it. I wanted to talk about the problem of some young children using credit cards a lot for online shopping without their parents' permission, But while I was struggling with English, the five minutes was over.

I: 5分間で最初3つ意見書こうと思ってて、意見だけ最初出したんですけど。後でDetailsを加えようと思って。(…)で3つ目の理由で最後に書ききれなかったんですけど、勝手にクレジットカードを登録していっぱいネットショッピングで使っちゃったよって言う問題を書こうと思って、そこで英語が出てこないってなったら5分終わってました。

During the 5-minute planning session, they could not write much more than topic sentences and three statements to support their positions for the outline. However, they added more details during the 12-minute timed-writing task. To reveal their writing process, I asked them to explain what they did to complete the 12-minute timed-writing task step by step. My interest concerned how much time was spent developing content while composing their text. Although the participants did not have time to write details on the outlining worksheet, they already knew what they wanted to write at the time of planning—in some cases, even before the planning session.

[Timing of an idea generation]

4. I: In the outline, you wrote “They don’t know well that” and the sentence ended here. In the composition written during the timed-writing task, you added more details. Did you already have these ideas during the planning stage or were they generated as you wrote the text?
P3: They were already in my mind during the five-minute planning time.

I: アウトラインは “they don’t know well that” で文が終了してるんですけど、timed-writingの方はもっと details が足されていますね。この部分っていうのは準備時間の時に既に頭の中にもありましたか、それとも書きながら出てきた感じですか。

P3: もう5分の準備の時点で頭の中にもありました。

P1: [In the reading passage] I guess there was a sentence “children come home late,” but that idea “[parents can use a smart phone] to contact their children.” was already in my mind as soon as I saw the topic.

P1: 子どもたちが遅くなるっていう文章が確かあったと思うんですけど、トピックをみた時に理由として「連絡する」っていうのは思いついてたんで。

These examples indicated that the participants were able to generate ideas without much support of peer-speech or the reading pre-tasks and by the end of the five-minute planning time, they already had ideas for their timed-writing before the oral rehearsal pre-task (Participant 3) and reading pre-task (Participant 1). Therefore, during the 12-minute timed-writing task, they focused on transcribing ideas into English rather than developing content. It was a stage where they needed to solve linguistic problems.

[Focusing on form]

5. I: So, your position was ‘disagree’ and then you wrote the first reason. How did you complete this part?

P3: I didn’t know how to use the word *manage*. [I asked myself] Do we say self-managing? Is it a noun? But I thought I could use the phrase *take care* instead and changed it.

I: At which phase did you wonder whether [the usage of the] *manage* could be wrong?

P3: During 12-minute timed-writing task. When I write, I read the text silently in my mind. And I began to wonder how I was supposed to use the word *manage*. Instead, I used *take care* more often, and it did not sound odd, so I changed it when I transcribed it into proper sentences.

I: 意見は disagree でまずは first reason ですね。ここはどんな感じでやりましたか。

P3: *Manage* の使い方がよくわからなくて、*self-managing* って言ったっけなあって思って、名詞かなあって思って。*Take care* の構文が使えるかと思ってそっちに変えたんですけど。

I: *Manage* が合ってるかなあって疑問に思ったのは、いつの段階ですか。

P3: 文を書いている 12 分の段階です。自分で文字を書くときは自分の頭の中でしゃべるんですけども、*manage* はどう使ったのかあやふやになっちゃって。どっちかって言うと *take care* を多用していたので、そっちだったらまだ違和感がないかなあと思ってちゃんとした文に直す段階で書き換えました。

I asked her when she noticed that *manage* could be wrong because in the outlining worksheet, she wrote “Elementary school students can’t *self-managing*,” but she changed the sentence to “Elementary school students can’t *take care of* themselves” in the text. She realized that her word choice did not sound correct as she recited the text while writing. The excerpts below are other examples of when another participant engaged in a focus on linguistic form during the 12-minute timed-writing task. After the oral rehearsal pre-task, she improved her utterances by using more appropriate and sophisticated lexis or multi-word expressions. She also worked on coherence.

[Polishing the oral output]

6. P2: Well...I changed the expressions of some sentences [while writing]. When I was speaking, sometimes I just uttered some words, not in complete sentences. In that case, I had to clean up the mess when I wrote to form proper sentences. For example, in the speech, I said “they have problem, their parents can learn danger”. [But in writing,] *even if* was added and “they have problems” was changed to “when problem occurs.”

P2: うーん。言い方を変えた文はありますね。やっぱり喋ってて、単語単語っていうか、文にすらなっていないことを話してる時があつて。そういうごちゃごちゃつてなった部分は多分書くときにうまく文になるよう、それこそ *even if* の後は話す時は *They have problem, their parents can learn danger* っとなつてたんですけど、*even if* つけたり、*they have problems* じゃなくて *problem occurs* にかえたりはしてます。

[Working on cohesion]

7. P2: When I spoke, I just uttered ideas in English one after another. But when I typed, I had to integrate all the utterances, so I added many words like *on the other hand* or *even if*, conjunctions not often used in spoken language, to connect each sentence.

P2: しゃべるときはその時その時頭にある考えを英語で1つ1つ出していくと思うんですけどタイピングするときに、話して英語になってる分を例えば *on the other hand* とか *even if* とか、接続詞とか喋る時あんま使わないんですけど、しゃべった内容をタイピングで打ち出すときに文がうまくつながるように結構付け足してあります。

In sum, the above examples indicate that the participants spent little time for idea generation during the five-minute planning period and the 12-minute timed-writing task because they could generate ideas on their own without relying on outside sources such as input from their partner or the reading passage. Instead, most of the time and cognitive resources were used for transcribing and monitoring. The topic of idea generation is discussed in relation to oral rehearsal and reading tasks in the following section.

To explore how the participants took advantage of each pre-task opportunity, I asked them which of the three pre-tasks they found most useful for improving fluency. Follow-up questions about the advantages and disadvantages of each writing pre-task were also asked.

Outlining

None of the interviewees responded that the outlining condition was the most effective; they did not have much to say about this pre-task. However, participant 1 said that even though this task was the least effective for improving writing fluency, she thought that the outlining task was useful.

[Outlining was cognitively demanding]

8. I: How did you feel about the outlining task?

P1: In the outlining task, I had to think of everything...grammar and everything by myself. It was useful, but when I wrote, I had nothing to refer to during the 12-minute timed-writing, and it was difficult to write.

I: You said it was “useful.” Could you explain a little more?

P1: In the outlining task, I did not get any information or grammar. Without anything to refer to, I had to think of many things, how should I say...I had to think about even details, so I had to use my head more.

I: Did you have to think about the content or language? Or both?

B: Both.

I: 最後の準備タスクのみはどうでしたか?

P1: やっぱり準備タスクだけだと自分で全部考えて文法も全部考えなきゃいけないし、その役には立つんですけど、文章を書くときに、その何も参考になるものがないとやっぱり 12 分間の時に、すごい書きづらいです。

I: 今「役に立つ」と言ったんですが、どういう点で役に立つんでしょうか?

P1: 5 分間のやつだけだと、誰からも情報っていうか文法とか、そういう参考になるものがない中でいろんなことをいっぱい考えようとするっていうか、なんでしょう、細かい部分まで考えなきゃいけないなくなるんで、その分結構頭を使ってる感じがします。

I: 考える内容っていうのは内容かな、それとも言語の方かな。それとも両方?

P1: どちらもです。

The excerpt points out that the outlining condition was cognitively demanding because the additional input provided in the two other conditions was unavailable. She used the word *useful* (役に立つ *yakuni tatsu*) to describe the outlining task because she had to think about many issue in detail. She thought it was good to use her head.

Oral rehearsal

According to the post-course questionnaire, the number of participants who selected oral rehearsal ($n = 15$; 39.5%) and reading ($n = 19$; 50.0%) were similar. Two interviewees said that the oral rehearsal pre-task task was the most effective. Participant 2 said that the oral rehearsal pre-task allowed her to produce language without translating from Japanese to English.

[Translating ideas directly to English]

8. P2: When I want to say something simple, English comes out naturally either in speaking or writing. But for long sentences or sentences with many words, I translate my ideas from Japanese to English first for writing. In speaking though, I must put some words in order [in my mind] and speak without stopping. So I can produce ideas more easily in English, without translating from Japanese to English.

P2: 簡単な英語だと考えてることがそのまま英語で[話し]言葉でもライティングでも出てくるんですけど、長めの文だったり複雑ないろんな単語が入ってくる文だと、一回日本語から英語にして頭で日本語にしてから英語に直すって言う作業がライティングにあって。けどスピーキングってどんどん単語を並べてしゃべるみたいな、意識してあまり止めないようにしゃべるので、日本語から英語に翻訳するっていう過程なく、そのまま考えていることを英語に出すっていうのがライティングよりは慣れている感覚があって。

Participant 3 said that the oral rehearsal task was most helpful because she was more used to speaking than reading and writing. In addition, she took advantage of the rehearsal opportunity.

[Taking advantage of the rehearsal opportunity]

9. P3: The speaking task was the most effective. I have received an English education that focused on speaking since my childhood, so I am very bad at reading and writing in English. Apart from these skills [reading and writing], I can speak fairly well in English, so I just said everything that came to my mind aloud and then all I had to do was to write down what I said. So I could write more with the speaking task.

P3: スピーキングがいちばん効果がありました。私小さい時から英語で喋る重視の英語教育を受けてきたので、英語の読み書きがすごく苦手なんです。それを抜きにしてしゃべるのだったら結構できるので、自分の頭の中で英語で思いついたことを口に出して、後はそれを口に出して文章に書き起こせばいいだけだったので、私としてはスピーキングがいちばんいっぱい書けました。

To examine whether the oral rehearsal task was useful for developing content or linguistic knowledge, I asked them whether they used the ideas or words they had heard in their partners' speeches.

[Using the expressions used in the partner's speech 1]

10. P3: I had a firm idea of what I wanted to say before the oral rehearsal task. But when I heard my friend's speech and felt "Oh, I didn't think that way," I stole it.

I: Was it your friend's ideas or expressions?

P3: Expressions.

I: Linguistically, do you often refer to your partner's speech?

P3: Yes. Mostly [I use it as reference] for linguistic expressions. (...). I was like "Ah-ha!" when my friend said what I wanted to say but did not know how to say it [in English]. I was like "Oh, I see!"

P3: 自分の言いたい事はもう話し合う前にはがっちり決まっているんですよ。ただ話を聞いたときにあーそういう視点があるのかっていうのがあると盗んじゃいます。

I: それはアイデアですか、それとも言い回しとか？

P3: 言い回しです。

I: 言語面でもパートナーのスピーチを参考にすることありますか？

P3: はいあります。ほとんどが言語面です。(...)なるほど！みたいな。こう言いたいんだけど自分じゃ言えないんだよなあって言う、なんていうんだろなあって言うのを向こうがしゃべったりしてると、そうかっ！てなります

[Using the expressions used in the partner's speech 2]

11. P2: When I didn't know how to turn my idea into English sentences and heard my friend saying what I wanted to say, I was like "Yes! That's exactly what I wanted to say." and stole the expressions or sentences.

I: When you refer to what your partner said, do you use the ideas or linguistic expressions? Or both?

P2: In my case, mostly linguistic expressions. When my partner said exactly what I wanted to say when I could not compile my ideas or I didn't know how I should say it in the oral rehearsal task, I was like "That's the one!"

P2: 考えはあるけどうまく文に言葉にできないってなったときに相手が喋るじゃないですか、それでそうそう私が言いたいことこれなんだよ！って言う時に結構表現方法だったり文だったり盗んだりします。

I: その相手の言ったことを参考にする時ってアイディアを参考にしますかそれとも言い回しとか言語表現の方を参考にしますか。それとも両方ありますか？

P2: 私は言語表現が多かったです。あんまり自分でまとまらないなあって言う時に、喋るときにどうやって言ったらいんだろうって言う時に相手が自分の言いたいことをぱっと言ってくれと、それ！ってなります。

I asked them if their partner's speech was useful for learning new words and expressions or developing content, or both. They both said that it was the words and expressions that they referred to in their partners' speeches but not the ideas. When their partner said something they wanted to say but could not say it in English, they expressed emotion with phrases such as "Ah-ha!" or "Yes! That's exactly what I wanted to say" and incorporated the words and expressions in their compositions. I asked participant 2 why she did not use her partner's ideas, and she said "It was not so difficult to think of three reasons [by myself] and I ran out of time just writing about my own ideas. I often learn new ideas [from the partner's speech] but I did not add those to my own writing due to the time constraints." (だいたい理由3つって埋まりやすいじゃないですか。自分の理

由書くだけで時間がいっぱいいっぱいになると思うので、確かにそうだなって事はあると思うんですけどそれを書き足したりとかはないですね。時間的に。) Participant 3 also said she did not need more content because she had too many ideas to write about. However, participant 1 said that as a linguistic reference, the passage written by the teacher was more reliable.

[The passage written by the teacher is more reliable]

12. I: How did you feel about the oral rehearsal task?

P1: Through the oral rehearsal task, I could learn the words or ideas I could not think of [by myself]. But sometimes I could not hear them clearly and I was not sure if my grammar was correct. So compared with reading task, I would rather use the oral rehearsal task as a reference for developing content.

I: Not really a good reference for words or grammar?

P1: And in terms of accuracy, the passage seemed to provide more details. Also, I could visually see [the text] and it is easier to understand.

I: スピーキングタスクはどうでしたか？

P1: スピーキングも課題文と同じように思いつかなかったような単語とか内容を聞くことができたんですけど、その聞き取れなかったりとか、あとは正しい文法かっていうのはわからないので。なのでこの課題文 [reading passage] よりかはどっちかって言うと内容が参考になるような感じですよ。

I: 語彙とか文法はそんなに参考にならない感じですかね。

P1: あと正確性で見るとやっぱりこの文章 [reading passage] の方が詳しい感じがするなあと思って。あと目で見ることができるし、わかりやすいっていうか。

Reading

One interviewee, participant 1, said that reading was most useful. She said that a passage was a helpful reference because it was linguistically more accurate than input

provided by her peers in the oral rehearsal task. She also indicated that the reading passages were a good resource for both learning lexical items and developing content.

[The passage was a reliable resource]

13. I: Of the three pre-tasks, which one did you find the most useful?

P1: The reading task. Making an outline after reading is, I could tell the language in the passage is more accurate than what my friend said [in the oral rehearsal task]. There were opinions for both agree and disagree, and I read the opinions of the person who was on the same side as me carefully. And then, I checked the words in the passage that I didn't know or could not recall by myself. Also, when I found an idea I could not think of by myself, I kept it in mind.

I: 3つのタスクの中で最も役に立ったのはどれですか？

P1: リーディングタスクです。文章を読んでから準備期間に入るっていうのは、文章自体が友達と言った文章よりも正確性がある正しい文章で書かれているのがわかるので。賛成反対が書いてあって、自分と同じ意見の文章を重点的に読むようにして、なんていうかわからなさそうな単語、おもいつかなそうな単語をチェックしておくようにして。あとは自分が思いつかないような発想とかがあったら、そこを参考にするようにして。

The excerpt below is from participant 1, and provides details for how the lexical item in the passage was incorporated in her writing. Because it was mentioned by the interviewees who did the oral rehearsal task, it is further evidence that shows that words were incorporated in writing tasks when they noticed what they wanted to say in the input.

[Using the lexical items from the passage]

14. T: You pointed out that you check the ideas and vocabulary in the passage while reading. Today, did you find any ideas or words in the passage useful for your writing?

P1: Well, I think there was the phrase *children are late* in the passage.

T: Late for what?

P1: Late to get home. As soon as I saw the topic, I thought I would say “[a smartphone can be used] to contact [their parents]. But I didn’t know how to say *osokunaru* (getting home late) in English. While reading, I found this phrase and I thought “Oh, I can use this.” I remember that clearly.

T: リーディングの内容とか単語とかを参考にしながら言っていたんですけど、例えば今日やった課題でリーディングを参考になった部分ってありましたか。

P1: 確か遅れる、子供たちが遅くなるっていう文章が確かあったと思うんですけど。

T: 何が遅くなった？

P1: 帰るのが。家に。トピックをみた時に理由として連絡するっていうのは思いついてたんですけど、その時に帰りが遅くなったって言う単語が思いついてなかったのを読んだときに出てきてあーこれ使えるなあって思って、結構それは覚えてます。

15. P1: When I saw the topic, I thought of three reasons and listed some related words like a word association game in my mind. When I found words I didn’t know in English in the passage...I can usually find them because the ideas in the passages are common, I check them very carefully.

P1: トピックを見た時に3つの理由を出して、その中で関連するワードを連想ゲームみたいな感じで自分の頭の中を出して、でこれ英語にできないっていうのを...大体まあ課題文とかでありがちなので載っていたりするんですけど、載ってた時にすごいチェックするようにしています。

In the above excerpt, participant 1 said that she found the ideas for the supporting points soon after she saw the topic and listed necessary words to express those ideas in her mind. The words she was looking for were often found in the reading passages and when she found the words she wanted to know, she checked them carefully and used them in her timed-writing.

I was also interested in whether the passage was useful for idea generations. In the interview, participant 1 indicated that the passage was a good resource for both obtaining linguistic information and developing content. However, further questioning revealed that the ideas she wrote for this particular composition were based on her past experience.

[Ideas are based on her past experiences]

16. I: How did you come up with the idea for the first reason?

S1: When I saw the topic, the word *the Internet* came to my mind. In my childhood, when I wanted to know something, I always asked my parents. But I couldn't ask them when they were not home. Now, children have smartphones and if I had one when I was an elementary school student, I could have found the answers [on my own] for my study. So reflecting on my childhood experience, the first reason came out.

I: 一番目の理由はどのように出てきましたか？

S1: スマートフォンって言うとやはりインターネットが出てきて。小さい頃わからないことがあったときにすごい両親とかに聞いてて。両親がいない時に聞けなかったから、今はスマートフォンがある時代で小学生の時にもし持ってたらって考えると、勉強とかでわからないことを調べられるなど思っで。少し自分の体験と重ね合わせて一番目の理由が出てきました。

She said that the other two supporting points were also based on her past experience.

Therefore, even though she said that reading was helpful for generating ideas in excerpt 13, the impact of the passage on timed-writing content was minimal for this particular topic. Participant 3 reported that she did not rely on a passage for idea generation.

[Already knew what she wanted to write]

17. I: Why did you think the reading task was the least effective?

P3: The content of the reading flowed out of my mind. It flowed out, so I was like "what was written in the passage?"

I: So, just simply making an outline was better?

P3: Yes.

I: Did you find that the reading task rather hindered writing?

P3: Rather...not necessary? Maybe?

I: Not necessary?

P3: Maybe not necessary. I felt like I was taking a test, taking an English test.

I: Was reading a passage a burden for you?

P3: I was overwhelmed by the extra information. So I often wrote about the things that were not relevant to the content of reading. Thus, I could manage with or without it.

I: The passage was not so helpful?

P3: It was not unhelpful, but I just did not use it.

I: リーディングは何故最下位だった？

P3: リーディングは読んでも内容が流れていっちゃうんですよ。外に出て行っちゃうんで、あれ、なんて書いてあったかなあって。

I: むしろそれよりはシンプルにアウトラインのみのほうがよかった？

P3: そうですね

I: リーディングがあるとかえって邪魔になる感じですか？

P3: というか…必要ない？…かなっていう感じ。

I: 別になくても？

P3: なくてもいいかなっていう。テストをやっている、英語のテストをやっている気持ちになりました。

I: かえって負担が大きくなっちゃう？

P3: プラスで情報がついていうと、おおって言うか。だからリーディングで読んだ内容と関係ないことないことも書いちゃってたんで、別にあってもなくてもって感じ。

I: リーディングはあんまり参考にならなかった？

P3: 参考にならなかったと言うよりはしなかった。

She said that reading was least helpful because she had a sufficient number of ideas to write about and additional resources were not necessary. I asked her further questions to understand why the outlining task was better than the reading task because in the reading task, she received more resources, 5 minutes for reading and an extra 5 minutes for planning. Participant 3 indicated that she was overwhelmed by the amount of information in the reading text and therefore did not remember much about the content.

Participant 2, who said that the oral rehearsal task was most helpful, however, commented that when she wrote about a topic she was not familiar with, the reading passage was useful.

[Passage was useful for unfamiliar topic]

18. I: Of the three pre-tasks, which one did you find most helpful?

P2: It's either oral rehearsal or reading. For some topics, I could write easily and others were difficult to write about. When writing about a difficult topic, I could use the ideas in the passage as a reference, and it helped me to write. Not considering that [topic familiarity], oral rehearsal task before timed-writing helped me to organize my ideas, so I would say that the oral rehearsal task was the best.

I: 3種類のタスクがあったんですけど、一番役に立つなと思ったのはどれですか。

P2: リーディングかスピーキングなんですけど。それこそお題によって自分が書きやすい内容と書きにくい内容ってあるじゃないですか。書きにくい内容の時ってリーディングがあるとそのリーディングの意見とかが参考になるので結構書きやすいんですけど、そういうのなしでいうとやっぱりスピーキングがあった方が私は書く内容がまとまりやすいのでスピーキングですかね。

To summarize, the interviews produced the following results.

1. The timed-writing topics were commonly used topics; thus, the interviewees did not have to rely on external resources such as input from a peer or a reading passage for idea generation or developing content. However, reading a passage might be useful for developing content when participants write about unfamiliar topics.
2. By the time the interviewees started writing, they generally knew what they wanted to express; thus, during the 12-minute timed-writing session, they mostly focused on producing their ideas in English. In this way, their cognitive resources were used to solve linguistic problems rather than developing ideas.
3. The words and expressions in a partner's speech and in the reading passages can be incorporated in a timed-writing text when the lexical items match what the writer wants to express. Therefore, oral rehearsal and reading pre-tasks can be good resources for solving lexical problems.
4. The outlining condition is challenging because it is necessary to generate ideas and form sentences without referring to external resources. However, this task can be useful because it is a cognitively demanding condition, and it induced language processing that the other two pre-tasks did not.

Additional Text Analysis of Lexical Items

The findings from the interviews indicated that reading passages can act as a useful resource for learning words and expressions. To examine the effect of the reading passage on the lexical diversity of the timed-writing, the word family similarity percentage between the passage and timed-writing was calculated for each pre-task group. The outlining and the oral rehearsal groups, those that did not read the passages, were also included for the comparison. Table 17 shows the results. The mean of the word family recycling index was 20.10, indicating that 20.10% of the word families in the timed-writing texts were shared with the reading passage. The means for the outlining and oral rehearsal groups were 19.23 and 20.19, respectively.

Although the outlining and oral rehearsal pre-task group did not read the passages, the percentages of the shared word families across the three groups were similar. This result indicates that most of the shared word families were high-frequency words and could be used by many participants without the support of the reading passage. Thus, although reading can serve as a useful resource for learning new lexical items, the reading passages did not significantly impact the selection of the words used in the timed-writing texts.

The Effect of Timed-Writing on Writing Development

To answer the third research question, which explored the longitudinal effect of the timed-writing task for one academic year on the development of L2 participants'

Table 17. *The Percentage of Word Family Similarity Rate Between the Reading Passage and Timed-Writing for Each Pre-Task Group*

	Outlining	Oral rehearsal	Reading
<i>M</i>	19.23	20.19	20.10
Median	19.10	20.15	20.27
<i>SD</i>	2.78	2.65	2.79
Minimum	12.77	14.17	14.86
Maximum	26.19	25.15	27.17
Skewness	0.34	-0.04	0.06
Kurtosis	0.58	-0.68	-0.10

Note. Word family similarity rate = the number of the shared word family / the number of word family in timed-writing + the number of the word family in the reading passages.

writing, a mixed design ANOVA was conducted separately for each of seven dependent variables—mean length of T-unit (MLT), dependent clause per clause (DC/C), mean length of clause (MLC), standard word units, lexical density (LD), moving-average type token ratio (50-word window) (MATTR), and measure of textual, lexical diversity (MTLD). The data was reviewed before the analysis, and two participants who missed more than two out of three tests were excluded from the data. Because the ANOVA was run seven times for each dependent variable, a Bonferroni procedure was used to control for Type I error. The alpha was set to .007 ($.05/7 = .007$).

Syntactic Complexity and Lexical Diversity

Mean Length of T-Unit

Table 18 shows the descriptive statistics for the mean length of T-unit. Throughout the academic year, from Time 1 to Time 3, the score of the comparison group was higher than that of the experimental group. The scores developed steadily for both groups as time progressed. In the case of the experimental group, the difference between Time 1 and Time 2 was subtle, but the growth between Time 2 and Time 3 was much greater (Figure 11).

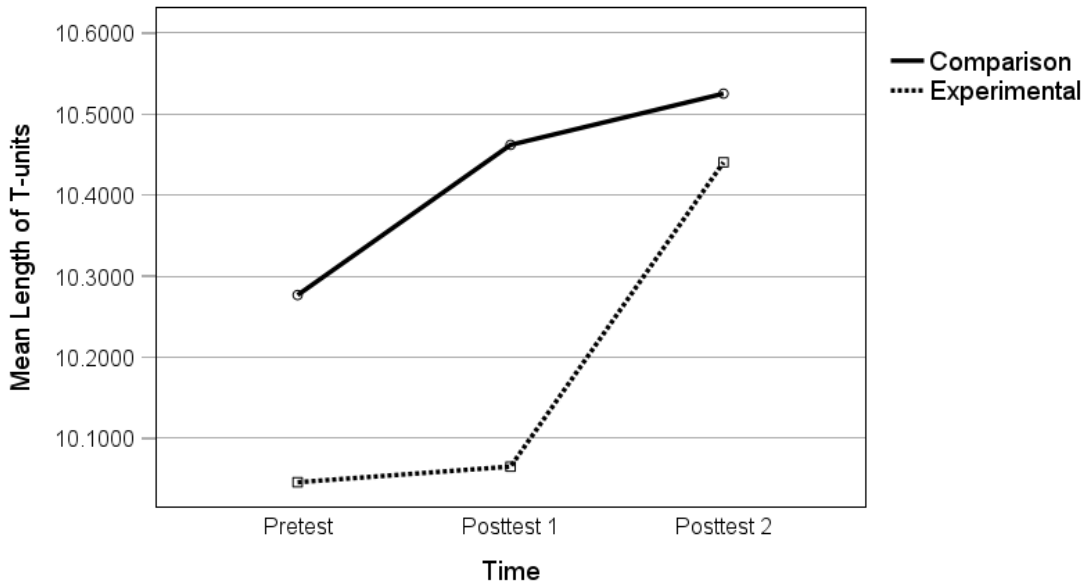
Before running the statistical tests, the z -skewness and z -kurtosis statistics were checked, and the normality assumption was met. After confirming the normality assumption, a mixed design ANOVA was conducted. The dependent variable was the score of the mean length of T-unit with three levels (the pretest, posttest 1, and posttest 2). The between-subjects factor was group with two levels, the experimental group and the comparison group. Mauchly's test of sphericity was significant, Chi-square = 10.01, $p = .007$. Thus, the Greenhouse-Geisser correction was adopted. There was no significant main effect of time, $F(1.79, 142.98) = .64, p = .51$, partial $\eta^2 = .008$, group, $F(1, 80) = .40, p = .53$, partial $\eta^2 = .005$, or interaction between time and group, $F(1.79, 142.98) = .14, p = .87$, partial $\eta^2 = .002$. The results showed that the mean length of T-units did not change over time and there was no difference between the experimental and comparison groups.

Table 18. *Descriptive Statistics for Mean Length of T-Unit for the Experimental and Comparison Groups*

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	10.05	10.07	10.44	10.28	10.46	10.53
<i>SD</i>	2.74	2.08	2.18	2.01	2.09	2.47
Skewness	1.97	0.94	1.29	0.74	0.86	1.62
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	6.13	0.32	2.85	-0.24	1.46	5.13
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group n -size = 40, Comparison group n -size = 42.

Figure 11. *Changes in the Mean Length of T-Units for the Experimental and Comparison Groups*



Dependent Clause per Clause

Table 19 shows the descriptive statistics for the dependent clause per clause measure. The experimental and comparison groups showed different trends. Although the score of the experimental group stayed almost the same throughout the intervention period, the scores of the comparison group decreased from Posttest 1 to Posttest 2 (Figure 12).

Before running the statistical tests, the z -skewness and z -kurtosis statistics were checked. It was determined that the normality assumption was met. After confirming the assumption, a mixed-design ANOVA was conducted. The dependent variable was the score of the dependent clause per clause with three levels (the pretest, posttest 1, and posttest 2). The between-subjects factor was groups with two levels, the experimental group and the comparison group. Mauchly's test of sphericity was not significant, Chi-

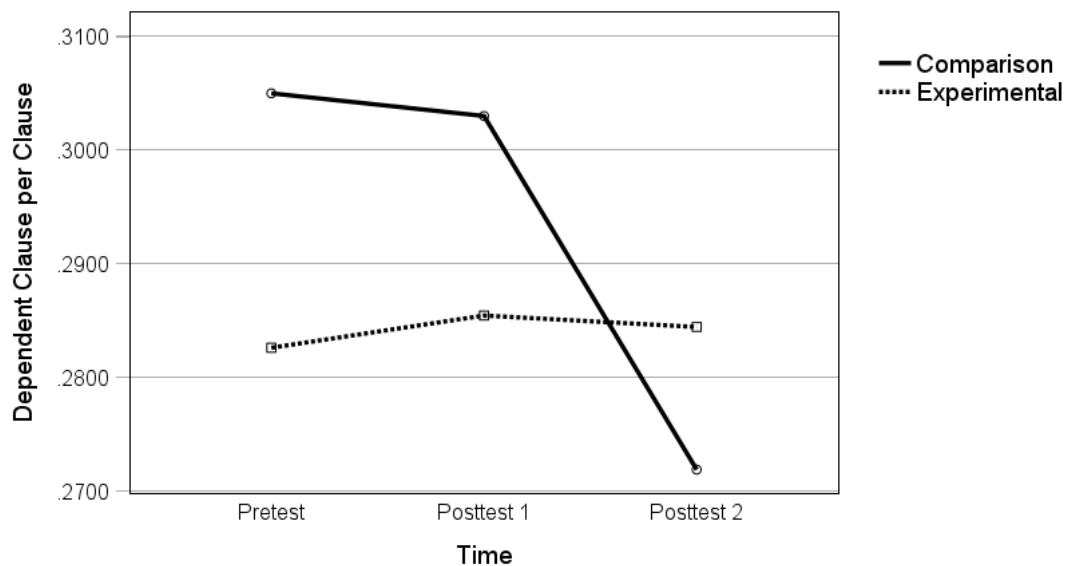
square = .20, $p = .90$. There was no significant effect of time, $F(2, 160) = .78, p = .46$, partial $\eta^2 = .01$, or group $F(1, 80) = .20, p = .66$, partial $\eta^2 = .002$, or interaction effect between time and group, $F(2, 160) = .84, p = .44$, partial $\eta^2 = .01$. The results showed that the score of dependent clause per clause did not change over time and there was no difference between the experimental and comparison groups.

Table 19. *Descriptive Statistics for Dependent Clause per Clause for the Experimental and Comparison Groups*

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	0.28	0.29	0.28	0.30	0.30	0.27
<i>SD</i>	0.14	0.11	0.11	0.11	0.12	0.12
Skewness	0.27	0.13	0.93	0.01	-0.11	0.39
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	0.13	-0.02	1.37	-0.32	-0.39	-0.82
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group n -size = 40, Comparison group n -size = 42.

Figure 12. *Changes in the Dependent Clause per Clause for the Experimental and Comparison Groups*



Mean Length of Clause

Table 20 shows the descriptive statistics for the mean length of clause. The experimental and comparison groups followed similar growth patterns. The scores improved as time progressed, and the difference between posttest 1 and posttest 2 was much greater than between the pretest and posttest 1 for both the experimental and comparison groups (Figure 13).

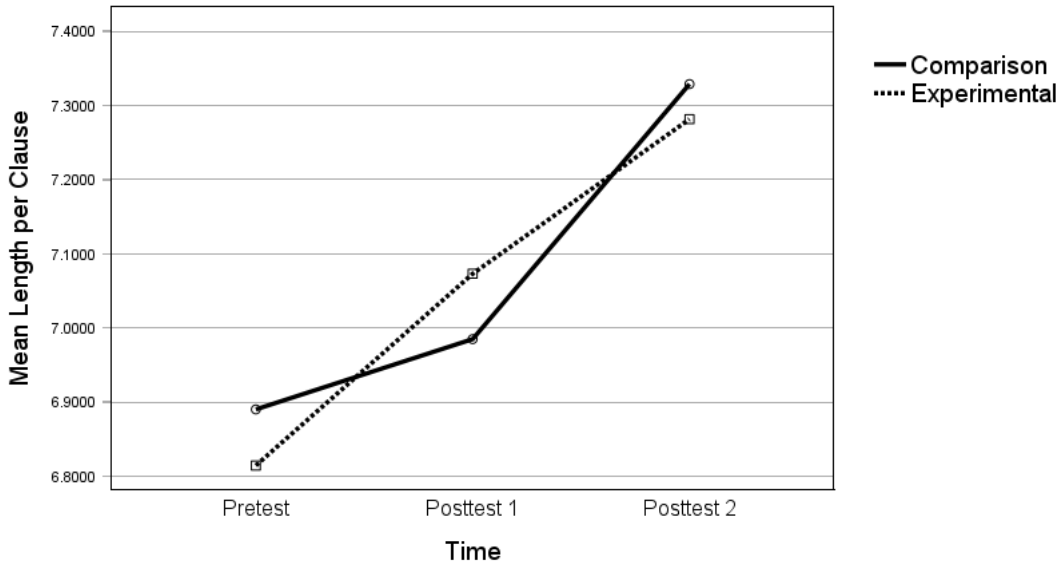
Before running the statistical tests, z -skewness and z -kurtosis scores were checked and the normality assumption was met. After confirming normality, a mixed-design ANOVA was conducted. Mauchly's test of sphericity was insignificant, Chi-square = 1.49, $p = .47$. The dependent variable was mean length of clause with three levels (the pretest, posttest 1, and posttest 2), and the between-subjects factor was group with two levels, the experimental group and the comparison group. The effect of time was significant, $F(2, 160) = 5.20, p = .006$, partial $\eta^2 = .06$. Contrasts revealed that there was a significant difference between the pretest and posttest 2, $F(1, 80) = 9.08, p = .003$, partial $\eta^2 = .10$. On the other hand, there was no significant effect for group, $F(1, 80) = .005, p = .94$, partial $\eta^2 = .000$, and no significant interaction between time and group, $F(2, 160) = .19, p = .83$, partial $\eta^2 = .002$. The results showed that the score of mean length of changed over time, but there was no significant difference between the experimental and comparison groups.

Table 20. Descriptive Statistics for Mean Length of Clause for the Experimental and Comparison Groups

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	6.81	7.07	7.28	6.89	6.98	7.33
<i>SD</i>	0.90	0.99	1.12	1.04	0.88	1.25
Skewness	0.00	0.88	1.40	0.90	0.09	0.80
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	-0.48	1.60	2.22	0.82	-0.05	0.14
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group *n*-size = 40, Comparison group *n*-size = 42.

Figure 13. Changes in the Mean Length of Clause for the Experimental and Comparison Groups



Writing Fluency: Standard Word Units

Table 21 shows the descriptive statistics for the standard word units. The figures indicated that both experimental and comparison groups improved as time progressed. Although the scores of the standard word units increased between the pretest and posttest 2 for both groups, the descriptive statistics showed that the difference between pretest and posttest 2 is greater for the experimental group than the comparison group.

Before running the statistical tests, z -skewness and z -kurtosis were checked, and normality assumption was met. After confirming normality assumption, a mixed design ANOVA was conducted. The dependent variable was the scores of the standard word unit with three levels (the pretest, posttest 1, posttest 2), and the between-subjects factor was group with two levels, the experimental group and the comparison group.

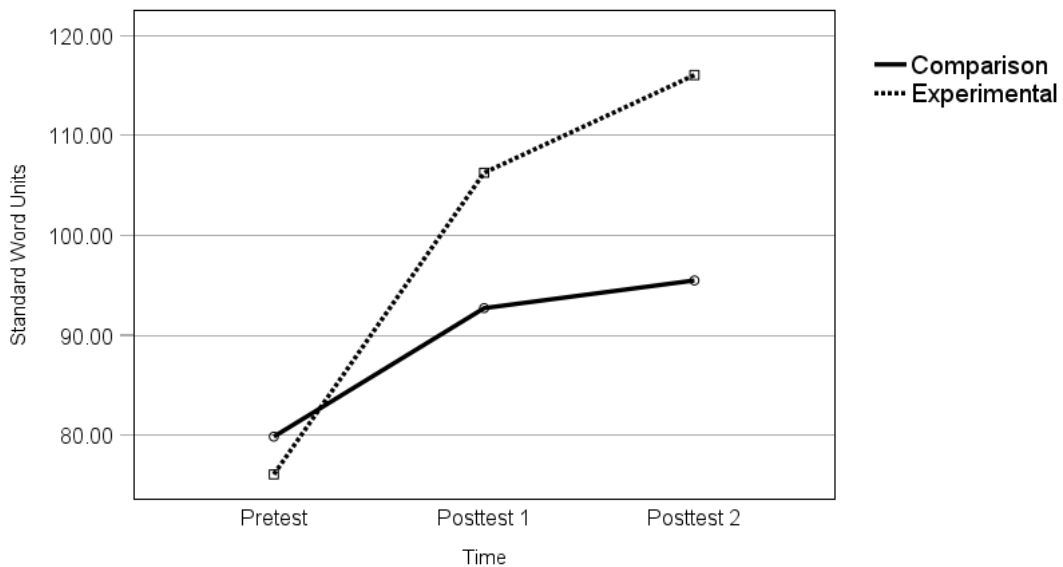
Mauchly's test of sphericity was not significant, $\text{Chi-square} = 4.03, p = .13$. The results showed there was a significant effect of time, $F(2, 160) = 40.97, p < .001$, partial $\eta^2 = .34$, group, $F(1, 80) = 3.97, p = .05$, partial $\eta^2 = .05$ and interaction between time and group, $F(2, 160) = 7.56, p < .001$, partial $\eta^2 = .09$. Contrasts revealed that there was a significant interaction of time and group between the pretest and posttest 1, $F(1, 80) = 7.46, p = .008$, partial $\eta^2 = .09$, and the pretest and posttest 2, $F(1, 80) = 17.33, p < .001$, partial $\eta^2 = .18$. This result indicated that the growth pattern of the standard word units between these periods differs between the experimental and comparison groups. Figure 14 shows the changes in the standard word units for the experimental and comparison groups. As the figure shows, the pretest score of the comparison group was higher than that of the experimental group, but the lines crossed after the pretest, indicating that the growth of the experimental group was greater than that of the comparison group.

Table 21. *Descriptive Statistics for Standard Word Unit for the Experimental and Comparison Groups*

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	76.08	106.26	116.04	79.86	92.72	95.49
<i>SD</i>	26.11	32.67	36.84	23.00	23.69	26.38
Skewness	1.61	0.51	0.07	1.06	0.70	1.80
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	5.72	0.34	-0.85	3.63	0.34	6.64
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group *n*-size = 40, Comparison group *n*-size = 42.

Figure 14. *Changes in the Standard Word Units for the Experimental and Comparison Groups*



Lexical Diversity

Lexical Density

Following the same procedures, the longitudinal effect of the timed-writing activity on lexical density was examined. Table 22 shows the descriptive statistics. The scores of both experimental and comparison groups increased from the pretest to posttest 1, but the two groups showed a different trend from posttest 2 to posttest 3. The

score of the experimental group decreased and that of the comparison group increased (Figure 15).

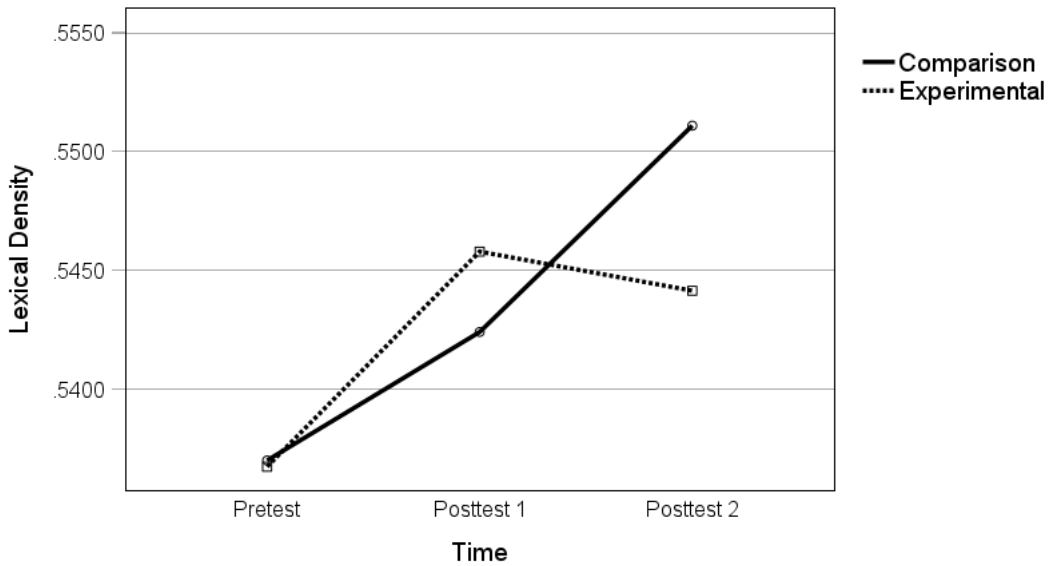
Before running the statistical tests, z -skewness and z -kurtosis were checked. After confirming the normality assumption, a mixed-design ANOVA was conducted. The dependent variable was the score of lexical density with three levels (the pretest, posttest 1, and posttest 2), and the between-subjects factor was groups with two levels, the experimental group and the comparison group. Mauchly's test of sphericity was insignificant, Chi-square = 3.47, $p = .18$. There was no significant effect of time, $F(2, 160) = 1.34, p = .27$, partial $\eta^2 = .02$, group, $F(1, 80) = .02, p = .88$, partial $\eta^2 = .000$, or interaction between time and group, $F(2, 160) = .31, p = .74$, partial $\eta^2 = .004$. The results showed that the lexical density score did not change over time and there was no difference between the experimental and comparison groups.

Table 22. *Descriptive Statistics for Lexical Density for Experimental and Comparison Groups*

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	0.54	0.55	0.54	0.54	0.54	0.55
<i>SD</i>	0.06	0.05	0.04	0.05	0.06	0.05
Skewness	-0.41	0.29	0.14	0.51	0.49	0.23
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	0.84	-0.32	0.07	1.06	-0.16	-0.40
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group n -size = 40, Comparison group n -size = 42.

Figure 15. *Changes in the Lexical Density for the Experimental and Comparison Groups*



Moving-Average Type Token Ratio

Table 23 shows the descriptive statistics for moving-average type token ratio (MATTR). The comparison group's score was higher than that of the experimental group throughout the academic year, but both groups had a similar growth pattern. The scores of both experimental and comparison groups decreased from the pretest to posttest 1, but they both recovered from posttest 1 to posttest 2 (Figure 16).

Before running the statistical tests, normality was assessed by checking the values of z -skewness and z -kurtosis, and the assumption was met. After confirming the normality assumption, a mixed-design ANOVA was conducted. Because Mauchly's test of sphericity was significant, $\text{Chi-square} = 12.62$, $p = .002$, the Greenhouse-Geisser correction was adopted. The dependent variable was the values of MATTR with three levels (the pretest, posttest 1, and posttest 2), and the between-subjects factor was group with two levels, the experimental group and the comparison group. There was no

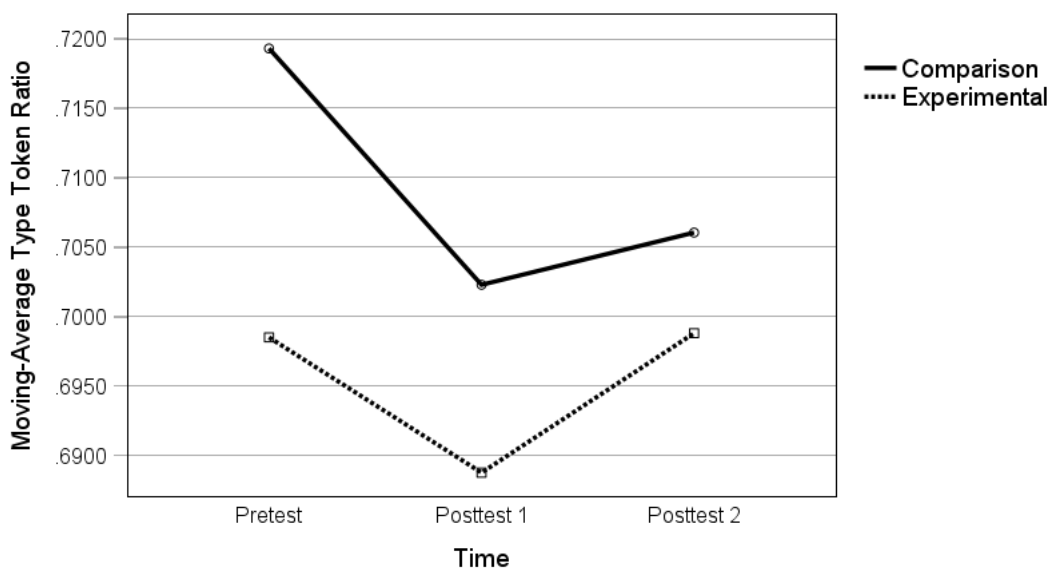
significant effect of time, $F(1.74, 139.41) = .71, p = .47$, partial $\eta^2 = .009$, group, $F(1, 80) = 1.24, p = .27$, partial $\eta^2 = .02$, or interaction between time and group, $F(1.74, 139.41) = .18, p = .80$, partial $\eta^2 = .002$. The results showed that the MATTR score did not change over time and there was no difference between the experimental and comparison groups.

Table 23. Descriptive Statistics for Moving-Average Type Token Ratio for the Experimental and Comparison Groups

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	0.70	0.69	0.70	0.72	0.70	0.71
<i>SD</i>	0.07	0.07	0.05	0.13	0.07	0.06
Skewness	-0.44	-1.09	-0.78	3.74	-0.11	-0.11
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	-0.06	1.06	1.48	20.18	0.20	0.03
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group *n*-size = 40, Comparison group *n*-size = 42.

Figure 16. Changes in the Moving-Average Type Token Ratio for the Experimental and Comparison Groups



Measure of Textual, Lexical Diversity

Table 24 shows the descriptive statistics for the Textual, Lexical Diversity (MTLD) measure. Different trends were observed for the experimental and comparison groups. The figure of the experimental group decreased from the pretest to posttest 1, but recovered on posttest 2. On the other hand, the comparison group’s score improved gradually over time (Figure 17).

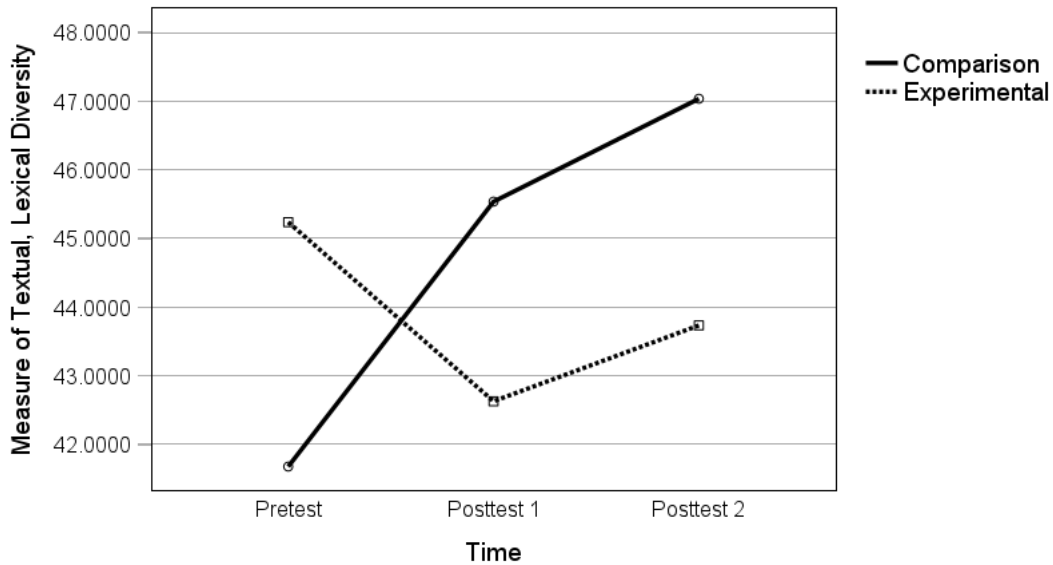
Before running the statistical tests, normality was assessed by checking the z -skewness and z -kurtosis, and the normality assumption was met. After confirming normality, a mixed-design ANOVA was conducted. The dependent variable was the MTLD scores with three levels (the pretest, posttest 1, posttest 2). The between-subjects factor was groups with two levels, the experimental group and the comparison group. Mauchly’s test of sphericity was insignificant, Chi-square = .50, $p = .78$. There was no significant effect of time, $F(2, 160) = .50, p = .61$, partial $\eta^2 = .006$, group, $F(1, 80) = .17, p = .68$, partial $\eta^2 = .002$, or interaction effect, $F(2, 160) = 1.92, p = .15$, partial $\eta^2 = .02$. The results showed that the MTLD score did not change over time and there was no difference between the experimental and comparison groups.

Table 24. *Descriptive Statistics for Measure of Textual, Lexical Diversity for the Experimental and Comparison Groups*

	Experimental group			Comparison group		
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
<i>M</i>	45.24	42.63	43.73	41.68	45.54	47.03
<i>SD</i>	15.51	12.34	11.53	11.86	15.01	17.51
Skewness	0.78	0.58	0.87	1.17	0.82	1.79
SES	0.37	0.37	0.37	0.37	0.37	0.37
Kurtosis	0.56	0.85	1.37	1.99	0.61	4.31
SEK	0.73	0.73	0.73	0.72	0.72	0.72

Note. Experimental group n -size = 40, Comparison group n -size = 42.

Figure 17. Changes in the Measure of Textual, Lexical Diversity for the Experimental and Comparison Groups



In sum, there was a significant effect of time for one syntactic complexity measure, mean length of clause. There was a significant difference between the pretest and posttest 2. There was significant interaction between time and group for a fluency measure, standard word unit, indicating that the experimental and comparison groups had different growth patterns. Although the scores of both groups increased, the gain of the experimental group was higher than that of the comparison group. This finding indicates that timed-writing was effective for improving writing fluency. There was no significant effect of time or group between the experimental and comparison groups for any of the three lexical diversity measures.

CHAPTER 6

DISCUSSION

In this chapter, I first review and interpret the results for the effect of the three pre-tasks on syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility. Next, I discuss the findings of the interviews and questionnaire, and additional text analyses. I then discuss the longitudinal effect of timed-writing on writing development, comparing the results with the comparison group. The time effect of the experimental group is discussed by incorporating the results of the pretest and two posttests (the macro-perspective) and four time points during the intervention period (the micro-perspective). Finally, I present research-oriented findings and the pedagogical implications based on the findings of this study.

The Effect of Three Pre-Tasks on Syntactic Complexity, Writing Fluency, Lexical Diversity, Content, and Comprehensibility

The first research question asked about the effect of three planning conditions—outlining, oral rehearsal, or reading—on syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility. In this section, I focus on the results of between-subject effects, and the effect of time is discussed in a later section with the pretest and posttest results.

Overall, the type of pre-tasks did not strongly impact syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility of the timed-writing task. There was, however, evidence in the interview and post-course questionnaire responses indicating that oral rehearsal or reading pre-tasks had a positive effect on

enhancing writing fluency, lexical diversity, and content of the timed-writing products, but the impact was not strong enough to cause statistically significant differences.

One possible reason is that the time spent for the pre-tasks and timed-writing was insufficient. In this study, the participants spent five minutes for planning, five minutes for the oral rehearsal and reading tasks, and 12 minutes for the timed-writing task. The findings from the interviews showed that the participants' priority was completing the timed-writing task. They had to focus on transcribing their ideas in English and had little or no time or cognitive resources left for selecting more complex grammar or vocabulary, or to develop the content of their timed-writing.

In relation to the first point, another explanation is the participants' English proficiency level. Johnson et al. (2012) indicated that for writers to take advantage of pre-task planning that reduces the working memory burden during the translating process, their language proficiency must reach a certain threshold. The participants in this study were low-proficiency learners who struggled to produce language. As the findings from the interviews showed, their priority was to complete the timed-writing task. Under time pressure, the participants had to concentrate on transcribing their ideas in English, and they did not have enough time or cognitive resources left to pay attention to linguistic aspects during the pre-task activities. In the following sections, I discuss this issue more in detail.

The type of pre-task did not strongly impact syntactic complexity. Syntactic complexity is a measure that is not readily affected in a short time by pre-tasks because the ability to form syntactically complex sentences develops gradually. The result

coincided with Johnson et al. (2012) and Nitta and Baba (2014), who found no significant effect of pre-tasks or task repetition on syntactic complexity in the short term. Previous literature indicated that pre-task activities can reduce the cognitive load of language processing (Kellogg, 1996). Reading a text, for example, can provide information concerning syntax that learners can use in their writing (Hirvela, 2016; Stotsky, 1995).

However, for low-proficiency English learners, applying syntax in their writing was too cognitively challenging. To incorporate more complex syntax in their writing, they must modify it into an appropriate form to fit their writing, and they must have the ability to use that particular grammatical form productively. If they were not developmentally ready, they would be unable to use that item in a situation involving time pressure, which is the case with time-writing tasks. Because they faced time constraints, it is understandable that their priority was to complete the timed-writing task, and they did not have enough cognitive resources available to pay attention to relatively complex syntax. That reason is a plausible explanation for why the pre-tasks did not significantly impact syntactic complexity on the timed-writing task.

The pre-task conditions also did not positively affect writing fluency. The result of this study coincides with Doe and Figueroa (2015), who found no statistical difference between writing planning (similar to outlining) and speaking planning (similar to oral rehearsal), and Johnson et al. (2012), who reported no significant effect for an outlining pre-task on writing fluency. The argument of Johnson et al. (2012) about the threshold of L2 proficiency can explain these results also. When the

participants' language proficiency has not developed sufficiently, the benefit of oral rehearsal is minimal. For example, observing the participants during oral rehearsal, I found that some of them were not ready to give their speeches and therefore their speeches were very short. In addition, the success of oral rehearsal can also be affected by one's partner. One interviewee commented that the input from peers is not always reliable and preferred the input provided by the reading passage written by the teacher. If a participant cannot fully take advantage of oral rehearsal or a partner cannot provide appropriate input, the oral rehearsal task might not reduce the cognitive load in working memory that is a necessary component in enhancing writing fluency.

As for the reading pre-task, some interviewees responded that they knew what they wanted to write without the support of a reading passage. This issue was mentioned by S1 in the interview (Excerpt 17). The time was limited, and her priority was to complete the timed-writing task rather than to develop content. Although the topic of the reading text was thematically the same as the writing topic, the genre of the writing was an opinion paragraph. Thus, the supporting ideas they wanted to use for their writing were not necessarily the same as the ones found in the reading passage. Unless the content of the passage matched what they wanted to say, they went with their own ideas, which might partly explain why the reading pre-task did not enhance writing fluency.

In the questionnaire, 50% of the participants selected reading as the most helpful pre-task for improving writing fluency because it helped them generate ideas, and 40% responded that the oral rehearsal pre-task was the most helpful. One interviewee

commented that the reading pre-task was useful for idea generation for unfamiliar topics (Excerpt 18). These comments indicated that some students benefited from the oral rehearsal and reading pre-tasks. On the other hand, as indicated above, there were cases where these pre-tasks were ineffective. Thus, the results were counterbalanced, and perhaps that finding explains why the pre-task effect was not significant.

The pre-task conditions did not affect the lexical diversity produced in the timed-writing task. One possible reason is that lexical diversity is a measure that does not change in a short period. Nitta and Baba (2014) indicated that in their study, fluency improvement occurred first, followed by syntactic complexity, and lexical diversity was the last to develop. Other studies have also shown that a significant development of lexical diversity was observed only for advanced (CEFF C1) students (Bulté & Housen, 2014; Mazgutova & Kormos, 2015). Therefore, even though the participants in this study were provided with input through their peers' speech or reading passages, they did not significantly impact the lexical diversity of the immediate timed-writing task.

Another possible reason is that more than 85% of the words in the passages were in the most frequent 1,000 words of English. Thus, there were not enough lower-frequency items to learn in the passages. The text analysis on lexical diversity also indicated that some words used in the reading passages were recycled in the compositions of the reading pre-task groups, but the similarity rates of the lexical items between the reading passage and compositions were not greatly different among the three pre-task groups, which shows that the participants used words they already knew.

These reasons can partly explain why the impact of the reading passage was not significant.

Although there was no statistically significant difference, the descriptive statistics showed that the MATTR and MTLT scores for the reading pre-task was the highest at all four time points. Thus, the reading task might have exerted a positive effect on lexical diversity. However, further research is necessary to investigate this possibility.

The type of pre-task did not affect the content of the timed-writing texts. In the questionnaire, the participants who selected the oral rehearsal or reading pre-tasks as the most helpful responded that those tasks helped them generate ideas. The survey results supported the descriptive statistics because the content scores for the oral rehearsal and reading pre-tasks was higher than that of the outlining group. However, the effect was not strong enough to produce statistically significant differences. Perhaps how much they could benefit from the reading passage depends on the topic. One interviewee responded that the reading pre-tasks were beneficial when they were writing about unfamiliar topics. The questionnaire was conducted at the end of the semester, so the participants' responses reflected their impressions of the entire intervention period. Those who responded that the tasks were useful for idea generation could have focused on the time when they had to write about difficult topics. Extra input might be helpful for developing content, especially with unfamiliar topics, but the impact was not strong enough result in a statistically significant difference.

In addition, in order to incorporate another individual's ideas in a written text, several steps are required. The writer must comprehend the input, select what they consider to be useful ideas, and paraphrase them in their own words, which consumes time and cognitive resources. Instead of using their limited cognitive resources for developing content, it is assumed that they decided to use it for transcribing, which is much more cognitively demanding and more important where task completion is concerned. The findings from the qualitative data also indicated that the linguistic expressions used by their peers or in the reading passages were useful and were incorporated into their timed-writing texts. The participants used their cognitive resources for dealing with what they found more important, and in the current study, transcribing was prioritized over developing content. Considering that timed-writing is a writing fluency task, their decision is reasonable and explains why the type of pre-task did not affect the content of the timed-writing products.

The type of pre-task did not affect the comprehensibility of the timed-writing texts. Although the focuses of the two studies is slightly different, the results of this study partially supported Zhang (2022). She reported that although the collaborative writing group (similar to the oral rehearsal pre-task in this study) outperformed the individual writing group (similar to the outlining pre-task) in terms of content and comprehensibility scores for a simple task, there was no significant difference in the scores between the two groups for a complex task.

In this study, the situation was perhaps similar to the case of the complex task in Zhang's study. The genre of the writing task used for this study is an opinion paragraph,

which requires inference skills and is considered more difficult than narrative or descriptive writing. Although the oral rehearsal and reading pre-tasks were expected to enhance comprehensibility, the participants were not fully able to benefit from these tasks because of the demand required for transcribing their ideas in English.

Finally, I would like to discuss the rubric used to evaluate content and comprehensibility. Another reason for the nonsignificant effect of the pre-tasks on content and comprehensibility might have been related to the rubric. The rating scale for functional adequacy created by Kuiken and Vedder (2016) was used in this study because they had validated the scale's reliability. However, with the participants in this study, the Rasch analysis indicated that the threshold distance between some adjacent categories was too close, especially between categories 3 and 4 (= 0.58) and between categories 5 and 6 (= 1.15). According to Linacre (1999), the difference in the adjacent threshold should be no less than 1.40., indicating that the distinctions between these categories were ambiguous and the scale included too many response choices. The participants in Kuiken and Vedder's (2016) study were Dutch L2 and Italian L2 learners whose CEFR levels were A2 to B1, while the majority of this study's participants were CEFR A2 Japanese L2 learners. The difference in the target language and proficiency level might explain why the scale did not function as well as expected.

Janssen et al. (2015) suggested reducing the number of categories in a scale when the threshold distance is too close. Thus, if a four or five-point scale had been used instead of a six-point scale, the rubric would have functioned better, and the results might have been different. Although there were no statistical differences, the descriptive

statistics showed that the reading and oral rehearsal pre-tasks outperformed the outlining pre-task for both content and comprehensibility scores. Thus, these two tasks might have some positive effect on content and comprehensibility, but more research in which an improved rubric is used is needed to investigate this possibility.

The Roles of Pre-Tasks

The second research question was an exploration of how the participants made use of each pre-writing task opportunity, and how it was reflected in the quality of their written products. To answer this question, I analyzed the post-course questionnaire, interviews, and compared the lexical items used in the reading passage and the timed-writing texts.

Outlining

Kellogg (1990) noted that producing a brief outline during pre-writing eases the cognitive burden of working memory during language processing and enhances writing fluency. However, the participants in this study did not benefit from outlining significantly. Because only four participants stated that the outlining pre-task was most helpful for timed-writing in the questionnaire, and some responses did not logically explain why it was helpful (e.g., “The quality of the outline will decide how easily I can write timed-writing”), it is difficult to interpret how this task did or did not help them in the timed-writing task.

The outlining pre-task was not chosen as the most helpful task because as mentioned by one interviewee, the learners had to think of everything from scratch without any external references to rely on. In this sense, this condition was the most

cognitively demanding. However, as Ortega (2005) reported, having extra planning time before the task could address the learners' lack of focus on syntactic structures. One interviewee said that although she found that the reading pre-task was the most helpful, the outlining task was also "useful" because it forced her to think deeply about the details of grammar and lexis. Descriptive statistics indicated that of the 12 cases (four time points x 3 measures), the scores for the outlining pre-task outperformed those of the other two pre-tasks. Although the participants did not perceive the outlining pre-task as helpful because it is cognitively demanding, it might be an effective way to enhance syntactic complexity. However, further research is needed to clarify this point.

Oral Rehearsal

The oral rehearsal pre-task was the second most popular pre-task after the reading pre-task. Some positive responses include that oral rehearsal was effective for idea generation ($n = 6$) and organizing ideas ($n = 6$). However, individual preferences for the oral rehearsal task can be influenced by individual difference variables and the success of pair-work. For example, Ortega (2005) reported there were accuracy-oriented and communication-oriented learners in her study. The former focuses on linguistic accuracy, while the latter prioritizes communication and is more tolerant of making mistakes. The differences in what they find important could affect their preference for particular types of pre-tasks.

Differences in individual preferences were observed among the participants in this study. For example, one interviewee [excerpt 9] said she had received a speaking-oriented English education and felt most comfortable with oral rehearsal pre-task. On

the other hand, another interviewee responded that reading was better because in terms of grammatical accuracy, the reading passages were more reliable than their partner's speech [excerpt 12] given that the former was produced by a native English speaker and the latter by a relatively low-proficiency second language user. Such individual differences in their interpretations of the task goals could have influenced their preference of the pre-task.

Furthermore, the presence of the partner could have affected the learners' task performance. One participant responded in the questionnaire that the oral rehearsal pre-task was not helpful for writing longer texts because when engaging in that pre-task, she used high-frequency words and simple grammatical structures and deleted some details so that her partner could easily understand her speech. However, for the reading pre-task, she could see English grammar structures in advance and express her ideas more easily.

Similar examples were observed in Ortega (2005), who argued that the presence of a partner can affect task performance. In her study, interview comments revealed that some participants told linguistically and propositionally simplified stories to their partners to make themselves more easily understood. The comments also showed that some learners hesitated to make self-corrections in order to avoid confusion. For the above reasons, the preference and success of the oral rehearsal pre-task was influenced by factors such as the participants' individual differences in learning style, their perceptions of the task goals, and the presence of the partner.

Reading

The reading pre-task had no statistically significant effect on syntactic complexity, writing fluency, lexical diversity, content, or comprehensibility. Notwithstanding this finding, 50% of the participants responded that the reading pre-task was the most helpful of the three pre-tasks. The most common reason was idea generation ($n = 16$), followed by vocabulary and grammar ($n = 8$). As previously mentioned, how much the participants could benefit from the reading pre-task partly depended on topic familiarity. The interview responses showed that it was good for them to know the opinions of both sides of the issue before writing. Thus, reading can be a particularly effective form of scaffolding when learners write about unfamiliar topics. In addition, the reading pre-task might enhance comprehensibility. The descriptive statistics indicated that the mean comprehensibility score for the reading pre-task was higher than that of the other two pre-tasks. The reading passage probably acted as a model, allowing the learners to visualize the form of the essay and providing them with useful target language input (Hirvela, 2016).

Regarding the effect of reading on lexical diversity, the word similarity rate between the reading passage and timed-writing texts written by the reading group was not significantly different from that of the other two groups because most of the words used in the passage were high-frequency lexis. Thus, as a whole, the effect of the reading passage on lexical diversity was small. However, some lower-frequency words in the reading passage were used in the reading group's timed-writing tests. For example, there were two occurrences of the word *jealous* in the writings of the reading

group. This word was used in the passage and was not observed in the texts written by the other two groups. Similar patterns were observed for the words *upload* (2 occurrences), and *survive* (2 occurrences). In addition, as mentioned in the previous section, the MATTR and MTLT scores for the reading pre-task outperformed those of the other two pre-tasks. As the evidence shows, the reading pre-task might be helpful for enhancing idea generation and comprehensibility, and can be the source for vocabulary. However, the result was not statistically significant, and further research is needed to clarify the effect of pre-reading on idea generation, comprehensibility and lexis in writing.

The Longitudinal Effect of Timed-Writing on Writing Development

The third research question asked how the experience of repeating the timed-writing task for one academic year contributed to the development of the participants' syntactic complexity, writing fluency, and lexical diversity. The timed-writing activity strongly impacted writing fluency. Both the experimental and comparison groups improved writing fluency over the academic year, but the growth of the experimental group was greater than that of the comparison group. There was a significant difference between the pretest and posttest for one of the three syntactic complexity measures, but there was no significant difference between the experimental and comparison groups. Regarding the lexical diversity measures, there was no effect for time or between-group differences for any of the three measures. In this section, I discuss the change by combining the results of two tests. One is the between-group comparison where the results of the pretest and two posttests for the experimental and comparison groups are

concerned. The other one is the development in writing over the academic year. The development of the experimental group is discussed by combining the two test results. One is a macro-perspective in which the scores for the pretest and two posttests were concerned. The other is a micro-perspective in which the scores of the four time points (Times 1–4) during the intervention period were compared.

There was a significant difference between the pretest and posttest 2 for mean length of clause (MLC) but no significant difference between the experimental and comparison groups, which indicates that both groups developed over the academic year. The fact that even the comparison group developed indicated that it was not just timed-writing activities that contributed to the growth. The participants in both groups had four 90-minutes English classes per week for two academic semesters and were exposed to a learning environment where they listened to and used English intensively. Thus, the growth of mean length of clause was probably due to the general English courses they took.

Looking at the development of mean length of clause and also mean length of T-units, which showed similar growth patterns from a micro-perspective, there was a significant difference between Time 1 and Time 2, and Time 1 and Time 3. However, after Time 3, the score decreased and there was no significant difference between Time 1 and Time 4. This result supports Bulté and Housen (2014), who suggested that in the short or medium term, complexity, accuracy, and fluency fluctuate rather than developing linearly. From the macro-perspective, there was development between the

pretest and posttest 2; however, over the one-year intervention period, the score fluctuated, as Bulté and Housen (2014) had found.

There was no statistical difference between the pretest and two posttests for two other syntactic complexity measures, mean length of T-unit (MLT) and dependent clause per clause (DC/C). Norris and Ortega (2009) indicated that the development of syntactic complexity first occurs with coordination, followed by subordination, and finally at the phrase level when the learners are advanced proficiency. In this study, however, MLC, which is categorized as a measure of phrasal complexity by Norris and Ortega (2009), developed most although the English proficiency of the participants was at the CEFR A2 level. This result partially supported Bulté and Housen (2014) who found significant changes in coordination and phrasal elaboration but not subordination. This result implies that although mean length clause was categorized as a phrasal complexity measure, it might be the one that develops at an early stage. Perhaps as Bulté and Housen argued, the development of syntactic complexity is not linear and does not strictly follow the pattern proposed by Norris and Ortega (2009). Further research is needed to clarify this point.

Both the experimental and comparison groups improved fluency over the course of the study; however, the interaction effect between time and group was significant for pretest and posttest 1, and pretest and posttest 2, indicating that the growth of the experimental group was more significant than that of the comparison group. From the micro-perspective, the differences between Time 1 and Time 2, Time 1 and Time 3, and Time 1 and 4 were all significant. Unlike the development of some syntactic complexity

measures that progressed slowly, writing fluency improved even in the early stages of the study, and it continued to improve throughout the academic year. The results indicated that the timed-writing activity led to improvements in writing fluency. Considering that timed-writing activity is a fluency activity, this result is not surprising. The students were told not to worry about mistakes and to prioritize quantity rather than quality in the timed-writing task. Each time the participants recorded the total number of words they wrote on the graph, they could visually see their progress. That might have helped them set goals and kept them motivated.

It is also possible to argue that writing fluency might have improved at the expense of syntactic complexity and lexical diversity. As Skehan (1998) proposed, L2 learners' working memory is limited, and it is not possible for them to focus on all three dimensions—fluency, syntax, and lexis—simultaneously. Thus, the participants in this study probably prioritized fluency. In addition, previous researchers have reported that fluency, syntactic complexity, lexical diversity do not develop simultaneously (Bulté & Housen, 2014; Mazgutova & Kormos, 2015; Nita & Baba, 2014). These studies have indicated that fluency develops first, followed by syntactic complexity, and finally lexical diversity. Because the participants in this study were at the CEFR A2 level, which is categorized as a basic language user, their developmental stage was probably at the phase of improving fluency; thus, the result can be seen as supporting previous studies. As suggested by Nita and Baba (2014), novice learners can better focus on formal aspects when they have developed a certain degree of fluency.

No significant change in the scores for any of the lexical diversity measures—lexical density, MATTR, and MTL—were observed between the pretest and the two posttests or between any time points during the intervention period. In addition, there was no significant difference between the experimental group and the comparison group. This result was consistent with Bulté and Housen (2014) and partially consistent with Nitta and Baba (2014), where a significant change in lexical diversity was reported for only the higher English proficiency group. Lexical diversity is an area that takes time to develop and general proficiency can be one factor that affects the differences in the results. In Bulté and Housen, the participants were at the CEFR B1 level and more growth was observed between the pretest and posttests for syntactic complexity than lexical complexity, whereas more growth was observed for lexical than syntactic indices for the C1 level participants in Mazgutova and Kormos (2015). Again, the hypothesis that the development of CFL measures occurs in order of fluency, syntactic complexity, and lexical diversity was supported. Most of the participants in the experimental group in this study were CEFR A2 level, and improving fluency was probably prioritized over syntactic and lexical aspects. The intervention period of this study was one academic year, but this period was still not long enough for development to occur in all three aspects of language: syntactic complexity, writing fluency, and lexical diversity. To examine the effect of fluency-focused writing activities on syntactic complexity and lexical diversity of low-proficiency learners, observing their development for more than one academic year is likely necessary.

Research-Oriented Findings

The findings of this study provide important implications for studies on writing pre-tasks and L2 writing development.

The first finding is that learners' general language proficiency and writing skills must be developed to a certain extent in order for them to fully benefit from writing pre-tasks. Although the design and the focus of Ruiz-Funes' (2015) study were slightly different from this study, the results of the current study were partially concurrent with those results. Ruiz-Funes' results showed different patterns for advanced learners and intermediate learners. Although high performers among the advanced learners achieved high scores for syntactic complexity, accuracy, and fluency for both simple and complex tasks, that happened on only the simple task for the lower performing advanced learners. A trade-off between syntactic complexity and fluency/accuracy occurred between the simple and complex tasks for the intermediate proficiency participants.

An interesting trend found in this study was that at the beginning of the intervention period, the descriptive statistics showed that the outlining pre-task outperformed the other two tasks in terms of syntactic complexity. The scores of all three syntactic complexity measures—MLT, DC/C, MLC—at Time 1 was the highest of the three pre-tasks. However, as time progressed, the scores of the oral rehearsal and reading pre-tasks groups improved steadily and at Time 4, the scores of all three syntactic complexity measures for outlining pre-task were the lowest. This trend implies that the participants were able to benefit from the oral rehearsal and reading task more

as they become more proficient and skilled at communicating in English. At first, simply making an outline worked better because the oral rehearsal and reading pre-tasks were too cognitively demanding. However, through continued engagement in the timed-writing task and enrollment in four 90-minute English classes per week, their writing skill and English proficiency improved. As a result, they were able to use their cognitive resources more efficiently and incorporate the language they used during the pre-task activities in their writing. The effect of the writing pre-tasks was not statistically significant in this study, but the oral rehearsal and reading pre-tasks might have a different impact on participants with higher English proficiency or an educational background in which they had more experience with communicative writing tasks.

Next, the findings of this study provided support for previous studies on the longitudinal effect of meaning-focused writing activities on writing development (Bluté & Housen, 2014; Nitta & Baba, 2014). As the results of these studies indicated, the development of CFL does not occur simultaneously, as each dimension follows a different developmental route. In this study, over the intervention period, writing fluency improved significantly, and some improvement was observed for one of the syntactic complexity measure, MLC, between Time 1 and Time 3. There was no difference between the pretest and two posttests or any of the four time periods for the lexical diversity measures. Thus, it seems that for L2 writing, fluency is the first dimension to develop, followed by syntactic complexity, and finally lexis. This result coincides with Bluté and Housen (2014) and Nitta and Baba (2014). Due to limited

cognitive capacity, foreign language learners first prioritize fluency and when they are fluent enough, they shift their attention to linguistic aspects such as grammar and lexis.

Finally, this study provided implications regarding the development of syntactic complexity. The trajectories of the mean length of T-unit, and mean length of clause followed a similar trend. They both increased gradually, but at Time 4, the scores for both measures decreased even though the scores of dependent clause per clause increased. This finding suggests that mean length of T-unit and mean length of clause are inversely proportionate to dependent clause per clause. One possible explanation is that as learners became more proficient, the occurrence of dependent clause per clause increased and as a result, the other two measures decrease. According to Norris and Ortega (2009), the development of syntactic complexity occurs in the order of coordination, subordination, and finally at the phrase level. In this study, there were more occurrences of mean length of clause, which is categorized as phrasal complexity, than dependent clause per clause, which is categorized as complexity in subordination at Time 2 and Time 3. Perhaps mean length clause develops at a relatively early stage of development. This study alone does not provide enough support for these implications, so further research is needed.

Pedagogical Implications

The findings of this study provide several pedagogical implications. First, although the impact was not large, each of three pre-tasks—outlining, oral rehearsal, and reading—can affect the performance of timed-writing task in different ways. By combining each pre-writing activity and a writing task effectively based on the

participants' needs and the educational situation, writing pre-tasks have the potential to enhance writing performances.

One of the findings of this study was that reading passages can provide useful L2 input to learners. Although there were no statistically significant findings, the descriptive statistics indicated that model reading passages can help learners retrieve receptive vocabulary items and lead to greater lexical variety. In addition, the descriptive statistics indicated that the reading pre-task has the potential for enhancing content and comprehensibility by providing a model for second language writers (Hirvela, 2016). These findings show that pre-reading can be an effective scaffolding activity for novice learners who need support for the transcribing stage. However, to make the task effective, learners should be provided with a model essay or a reading passage that is thematically the same as the writing topic and appropriate for their language proficiency level. By connecting reading and writing tasks effectively, it is possible to enhance learners' writing performances.

The oral rehearsal pre-task also provides advantages. First, although the result was not statistically significant, the oral rehearsal pre-task has the potential for improving writing fluency. Of the three tasks, the fluency scores for the oral rehearsal group for Time 3 and Time 4 were the highest. In pair-work, they must tell their partners their opinions. This pre-task is productive, it gives the learners reasons to communicate, and it puts them under pressure to produce utterances that can then be used in the writing task.

Furthermore, the interview and survey results revealed that input from peers is also useful for writers. The advantage of the input from peers is that it is provided by someone with a similar English proficiency level. The lexis used in peer interaction might be noticed relatively easily and incorporated in writing tasks. Peer input might contain non-target-like forms, but learners can benefit from it in a way that differs from teacher input. The descriptive statistics indicated that the score of one of the lexical diversity indices, lexical density for oral rehearsal, outperformed the other two pre-tasks. This finding provides initial evidence that learners can benefit from interactions with peers. Although both interlocutors are English-language learners, interactions with peers provides opportunities for comprehensible input and output.

Although there are possible advantages for reading and oral rehearsal pre-tasks, the findings of this study also indicated that integrated tasks can be challenging for low-proficiency learners. The effect of the pre-tasks on syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility failed to reach statistical significance. The oral rehearsal task can be especially cognitively demanding when learners are not proficient enough, as indicated by the comments from the interview and questionnaire. The results of this study showed that at Time 1, the outlining pre-task resulted in higher performances than the other two pre-tasks. Thus, at the beginning of the semester or when teaching novice learners, the outlining task might be a good one to start with, as learners would be able to benefit from the oral rehearsal and reading pre-tasks more when they become more proficient and more skilled at writing.

Next, I present the implications of using timed-writing activities in L2 classrooms. Continuing timed-writing longitudinally can positively affect learners' writing fluency. Extensive practice of the timed-writing task provided the participants with the opportunity to produce output and allowed them to write longer texts in the given amount of time. In EFL environments where the chance of using the L2 is limited, implementing a meaning-focused writing activity in each lesson can promote writing development. Continued practice is important to retaining growth because the fluency scores of many participants decreased after six weeks of summer vacation and then they recovered after two to three weeks of instruction.

The findings also revealed that it is important to know the developmental stage of the three dimensions of language investigated in this study—syntactic complexity, writing fluency, and lexical diversity. Fluency is the first dimension to develop. Therefore, timed-writing activity is an effective activity for beginner learners to improve their fluency. When learners are fluent enough, they can shift their attention more to formal aspects of the language such as syntax or lexis. This developmental order indicates that lexis is the last dimension of the language to develop and substantial changes might not occur until the learners' language proficiency level reaches a particular stage. To encourage intermediate and upper intermediate learners to attend to lexis, one possibly useful activity is to have them use an automatic lexical analyzer to analyze their own written text. For instance, the vocabulary profiler provided by Cobb (2023a) provides participants with data such as the number of word families or the percentage of words in each word frequency level. Through this activity, the

participants can be more aware of the words they have used in their writing and attempt to use a greater variety of lexis in later drafts or writing tasks.

CHAPTER 7

CONCLUSION

In this chapter, I first summarize the main findings of the study. I then discuss the limitations of the study and make suggestions for future studies. Finally, I close the chapter with concluding comments.

Summary of the Findings

To investigate the effect of three writing pre-tasks on writing quality and the longitudinal effect of implementing meaning-focused timed-writing activities on the development of syntactic complexity, fluency, and lexical diversity, this study was conducted for one academic year at a Japanese university. The study produced three main findings.

Regarding the effect of pre-tasks, none of the three pre-tasks impacted syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility measures to a statistically significant degree. The participants prioritized task completion or transcribing their ideas into English sentences, and contrary to my expectations, the oral rehearsal pre-task and reading pre-task did not significantly reduce the cognitive burden of transcribing. It is assumed that the participants could not fully take advantage of the pre-task because of their relatively low English proficiency level. Because the participants were CEFR A2 level learners, integrating speaking or reading pre-tasks with the time-writing task might have been too challenging. However, the descriptive statistics indicated a tendency for the oral rehearsal pre-task to lead to greater writing

fluency and content, and for the reading pre-tasks to lead to improvements in lexical diversity, content, and comprehensibility.

Regarding the role of each type of writing pre-task activity and how they affected the final written products, the results of the questionnaire and interviews indicated that nearly 50% of the participants found the reading pre-task most helpful. They responded that it was effective in terms of idea generation and learning vocabulary and grammar. The descriptive statistics indicated that MATTR and MTLTD scores for the reading pre-task outperformed those of the oral rehearsal and outlining pre-tasks. In addition, the additional text analysis on lexical items revealed that some low-frequency words in the reading passages were used only by the reading pre-task group. Although the impact was not strong enough to reach statistical significance, the analysis indicated that model reading passages can provide useful linguistic input for writing tasks. Future study is necessary to clarify this issue.

The oral rehearsal pre-task was the second most popular and about 40% of the participants found it the most helpful of the three pre-tasks. Some positive comments on the questionnaire indicated that it was effective for idea generation and organizing ideas. However, one student responded in the questionnaire that she reduced the information and used simpler language in her speech so that her partner would understand her easily. Others responded that the reading input was more reliable in terms of the accuracy of word usage and grammar than the input from a classmate. One commented that sometimes she had to start speaking before she was ready. Thus, how

much learners can benefit from the oral rehearsal pre-task is more influenced by individual differences compared with the other two pre-tasks.

Although some participants found the outlining task useful, only 10% of the participants perceived it as the most effective. When engaging in the outlining pre-task, the participants could not rely on any external resources for idea generation and transcribing their ideas in English. However, the descriptive statistics indicated that the outlining condition outperformed the other two pre-tasks in terms of syntactic complexity at some time points, especially in the early phase of the intervention period. There was also evidence from the interview that the participants focused their attention on formal aspect of the language during the outlining pre-task. Thus, making an outline might positively affect syntactic complexity, but this study could not provide enough evidence to support that possibility.

Lastly, regarding the longitudinal effect of implementing the timed-writing activity, the current study showed that it had a significant effect on developing writing fluency. After continuing timed-writing activities for one academic year, greater gains were observed for the experimental group than for the comparison group. For syntactic complexity and lexical diversity, there were no significant differences between the pretest and two posttests for either group except for mean length clause. However, the growth was significant for both the experimental group and the comparison group, indicating that it was not merely the effect of timed-writing.

Limitations of the Study

This study has three main limitations. First, the time allocated for the pre-tasks might have been insufficient. The participants spent only five minutes for the outlining pre-task, ten minutes for oral rehearsal and reading pre-tasks, and 12 minutes for timed-writing. However, time on task was a difficult factor to control because this study was conducted as a part of normal class activities, and I could not spend more class time for these activities. In addition, I thought that giving the participants too much time would not positively affect their performances because some participants get exhausted and cannot fully use the time to complete the task. Previous studies have indicated that prolonged pre-task planning and online planning do not necessarily result in the best performances (Ong, 2010; Rostamian et al., 2018). Thus, the results might not change significantly even if more time had been given to the participants.

Second, I could have used a better rubric to evaluate content and comprehensibility. I used the scale developed by Kuiken and Vedder (2017), the reliability of which was supported by their study. However, the FACETS analysis indicated that the distance between some categories, especially between Category 3 and 4 was too close. The pre-tasks did not have a significant effect on content and comprehensibility of the timed-writing task in this study, but the results might have differed if the validity of the scale had been examined and adjusted for the participants of this study.

The last limitation is a student and teacher factor. The participants were university students taking several English courses other than the course in which this

study was conducted. I could not control the influence of the teachers and English activities completed outside of this course. In addition, the experimental group and comparison group were taught by different instructors. There might have been some activities in the comparison group that impacted the results. Moreover, the English proficiency of the participants in each group was not equivalent. Thus, the comparisons of the experimental group and comparison group must be interpreted cautiously.

Suggestions for Future Research

Based on the findings, three suggestions can be made for future research. First, an improved rubric can be developed to evaluate content and comprehensibility. As indicated in the Limitations section, the distance between some categories were too close. One possible solution is to use a five-point scale instead of a six-point scale because the occurrence of Category 1 was extremely small. Another solution is to separate two evaluation points, adequacy and consistency, included in the content rubric. The distinction between some categories was unclear probably because there were cases when these two items could not be judged using one criterion. For example, compositions with an adequate number of ideas can be inconsistent.

Second, descriptive statistics indicated that the certain pre-tasks positively affected the syntactic complexity, writing fluency, lexical diversity, content, and comprehensibility scores. These scores might be statistically significant with a different research design and a larger number of participants. A Latin-squares design was adopted for this study, so the participants had opportunities to engage in all three types of pre-tasks. However, if the same intervention is assigned to each class throughout the

semester, the impact of the pre-task might have been different. Another suggestion about the design is providing a “no planning” group. In this study, the oral rehearsal and reading pre-task groups also completed the outlining task; therefore, the effect of outlining alone was ambiguous. If a no planning group is included, the effect of the outlining pre-task would be clearer.

The writing genre is another factor that might have impacted the results. An opinion paragraph was used for this study, but the impact of the reading pre-task might be larger for a genre such as storytelling. In the case of the opinion paragraph, what the participants wanted to say and the content of the passage might not always match. On the other hand, it is easier to control the content of a story telling writing task, and the participants can borrow ideas and linguistic expressions used in the passage more readily. By changing the writing genre, the impact of the reading pre-task might be stronger.

Finally, it would be useful to conduct a study that runs for more than one academic year to investigate how syntactic complexity, writing fluency, and lexical diversity measures continue to develop in EFL environments. Foreign language proficiency develops slowly and CFL measures do not develop simultaneously. In the case of the participants in this study, whose English proficiency level was CEFR A2, their growth in fluency was largest over the intervention period. Observing how other aspects such as syntactic complexity or lexical diversity develop for more than one academic year would provide additional insights to the research on L2 writing development.

Concluding Comments

Improving writing skills is a complex task and for foreign language learners who have few chances to receive input and use the foreign language, it is a challenging skill to acquire. In such an environment, meaning-focused writing activities such as timed-writing are a good way to practice the target language. The timed-writing task is especially effective for developing writing fluency for low-proficiency language learners, as it gives them opportunities for producing output, which is necessary for language development.

Traditionally, English writing instruction has strongly emphasized linguistic accuracy, especially in secondary schools in Japan, and little attention has been paid to improving fluency. However, by frequently engaging in fluency-based writing activities, students can develop greater fluency and focus their attention on linguistic aspects without explicit feedback from the teacher. Moreover, by combining fluency-based writing activities with pre-tasks, instructors can provide learners with greater opportunities to acquire new lexis and enhance the content and comprehensibility of their writing. I hope that this study provides useful input for teachers making lesson plans for writing classes and leads to more research on task-based L2 writing and writing development.

REFERENCES

- Abrams, Z., & Byrd, D. R. (2016). The effects of meaning-focused pre-tasks on beginning-level L2 writing in German: An exploratory study. *Language Teaching Research*, 21(4), 434-453. <https://doi.org/10.1177/1362168815627383>
- Amiryousefi, M. (2016). The differential effects of two types of task repetition on the complexity, accuracy, and fluency in computer-mediated L2 written production: A focus on computer anxiety. *Computer Assisted Language Learning*, 29(5), 1052–1068. <https://doi.org/10.1080/09588221.2016.1170040>
- Baba, K. & Nitta, R. (2010). Dynamic effects of task type practice on Japanese EFL students' writing: Text analysis with Coh-Metrix. *Proceedings of the Twenty-Third International Florida Artificial Intelligence Research Society Conference. FLAIRS*.
- Baddeley, A. D. (1986). *Working memory*. Oxford University Press.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Erlbaum.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. <https://doi.org/10.1016/j.jslw.2014.09.005>
- Bygate, M. (1996). Effect of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.). *Challenge and change in language teaching* (pp. 134–146). Routledge.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Routledge.
- Calkins, L. M. (1986). *The art of teaching writing*. Heinemann.
- Chen, Y. S., & Lin, M. F. (2021). Effects of peer collaboration on EFL learners' comprehension of conversational implicatures. *System*, 97, 1–12. <https://doi.org/10.1016/j.system.2020.102441>

- Cho, M. (2019). The effects of prompts on L2 writing performance and engagement. *Foreign Language Annals*, 52(3), 576–594. <https://doi.org/10.1111/flan.12411>
- Cobb, T. (2023a). Compleat Web VP (Version 2.6). [computer program]. <https://www.lex tutor.ca/vp/comp/>
- Cobb, T. (2023b). Text Lex Compare (Version 6). [computer program]. https://www.lex tutor.ca/cgi-bin/tl_compare/
- Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type-token ratio (MATTR). *Journal of Quantitative Linguistics*, 17(2), 94–100. <https://doi.org/10.1080/09296171003643098>
- Craver, R. P. (1976). Word length, prose difficulty, and reading rates. *Journal of Literacy Research*, 8, 193–203. <https://doi.org/10.1080/10862967609547176>
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Sage.
- Delaney, Y. A. (2008). Investigating the reading-to-write construct. *Journal of English for Academic Purposes*, 7(3), 140–150. <https://doi.org/10.1016/j.jeap.2008.04.001>
- Dobao, A. F. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, 21(1), 40–58. <https://doi.org/10.1016/j.jslw.2011.12.002>
- Doe, T., & Figueroa, A. (2015). Planning activities and freewriting. In P. Clements, A. Krause, & H. Brown (Eds.), *Focus on the learner* (pp. 405–411). JALT.
- Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 4–34). Benjamins.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474–509. <https://doi.org/10.1093/applin/amp042>
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59–84. <https://doi.org/10.1017/S0272263104026130>

- Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Eds.), *Planning and task performance in a second language* (pp. 167–192). Benjamins.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245. <https://doi.org/10.1037/0033-295X.102.2.211>
- Esmaeili, H. (2002). Integrated reading and writing tasks and ESL students' reading and writing performance in an English language test. *Canadian Modern Language Review*, *58*(4), 599–620. <https://doi.org/10.3138/cmlr.58.4.599>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage.
- Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, *31*(1), 21–32. <https://doi.org/10.2307/356630>
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387. <https://doi.org/10.2307/356600>
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*, 299–323. <https://doi.org/10.1017/S0272263100015047>
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, *3*, 215–247. <https://dx.doi.org/10.1177/136216889900300303>
- Gillian, F. & Sakamoto, M. (2008). *For and against*. Seibido.
- Gass, S., Mackay, A., Alvarez-Torres, M., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, *49*, 549–580. <https://doi.org/10.1111/0023-8333.00102>
- Google Classroom [Online class management tool]. (2023). <https://edu.google.com/workspace-for-education/classroom/>
- Good Calculators [Computer program]. (2023). <https://goodcalculators.com/flesch-kincaid-calculator/>
- Hirvela, A. R. (2016). (2nd ed.). Source-based writing: Reading to write and writing to read. In D. Belcher & J. Liu (Eds.), *Connecting reading & writing in second language writing instruction*. University of Michigan Press.

- Kessler, M., Ma, W., & Solheim, I. (2022). The effects of topic familiarity on text quality, complexity, accuracy, and fluency: A conceptual replication. *TESOL Quarterly*, 56(4), 1163–1190. <https://doi.org/10.1002/tesq.3096>
- Kim, K. J., & Pae, T. I. (2021). Examining the simultaneous effects of L1 writing, L2 reading, L2 proficiency, and affective factors on different task types of L2 writing. *Discourse Processes*, 58(7), 662–680. <https://doi.org/10.1080/0163853X.2021.1872989>
- Janssen, G., Meier, V., & Trace, J. (2015). Building a better rubric: Mixed methods rubric revision. *Assessing Writing*, 26, 51–66.
- Japan Statistical Society (2012). *Deta no bunseki* [Analyzing Data]. Nihon Toukei Gakkai.
- Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, 21(3), 264–282. <https://doi.org/10.1016/j.jslw.2012.05.011>
- Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, 37, 13–38. <https://doi.org/10.1016/j.jslw.2017.06.001>
- Kang, S., & Lee, J. H. (2019). Are two heads always better than one? The effects of collaborative planning on L2 writing in relation to task complexity. *Journal of Second Language Writing*, 45, 61–72. <https://doi.org/10.1016/j.jslw.2019.05.002>
- Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 143–164). Benjamins.
- Kellogg, R. (1990). Effectiveness of prewriting strategies as a function of task demands. *The American Journal of Psychology*, 103(3), 327–342. <https://www.jstor.org/stable/1423213>
- Kellogg, R. (1996). A model of working memory in writing. In C. Levy & S. Ransdell (Eds.), *The science of writing* (pp. 57–71). Erlbaum.

- Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* [Doctoral dissertation, Georgia State University].
<https://doi.org/10.57709/8501051>
- Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition*, 43, 781–812.
<https://doi.org/10.1017/S0272263120000546>
- Kyle, K. (2023). Tool for the automatic analysis of syntactic sophistication and complexity (Version 1.3.8). [computer program].
<https://www.linguisticanalysisistools.org/taassc.html>
- Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal*, 102, 333–349. <https://doi.org/10.1111/modl.12468>
- Kyle, K., Crossley, S. A., & Jarvis, S. (2020). Assessing the validity of lexical diversity using direct judgements. *Language Assessment Quarterly*, 18(2), 154–170.
<https://doi.org/10.1080/15434303.2020.1844205>
- Kyle, K., Crossley, S. A., & Jarvis (2023). Tool for the automatic analysis of lexical diversity (Version 1.4.1). <https://www.linguisticanalysisistools.org/taaled.html> [computer program].
- Kuiken, F., & Vedder, I. (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing*, 34(3), 321–336.
<https://doi.org/10.1177/0265532216663991>
- Lapkin, S., & Swain, M. (2000). Task outcomes: A focus on immersion students' use of pronominal verbs in their writing. *Canadian Journal of Applied Linguistics (CJAL)*, 3, 7-22. <https://journals.lib.unb.ca/index.php/CJAL/article/view/19821>
- Lapkin, S., Swain, M., & Smith, M. (2002). Reformulation and the learning of French pronominal verbs in a Canadian French immersion context. *The Modern Language Journal*, 86(4), 485–507. <https://www.jstor.org/stable/1192721>
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. The MIT Press.
- Liao, M., & Wong, C. (2010). Effects of dialogue journals on L2 students' writing fluency, reflections, anxiety, and motivation. *Reflections on English Language Teaching*, 9(2), 139–170. Retrieved from
<http://www.nus.edu.sg/celc/research/books/relt/vol9/no2/0-prelim.pdf>

- Linacre, J. M. (1999). Investigating rating score category utility. *Journal of Outcome Measurement*, 3(2), 103–122. Retrieved from http://jampress.org/JOM_V3N2.pdf
- Linacre, J. M. (2023). Facets computer program for many-facet Rasch measurement (version 3.85.1). <https://www.winsteps.com/facets.htm>
- Linacre, J.M. (2023). A users guide for FACETS. <https://www.winsteps.com/a/Facets-Manual.pdf>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45(1), 36–62. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190–208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x
- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, 29, 3–15. <http://dx.doi.org/10.1016/j.jslw.2015.06.004>
- McCarthy, P. M. (2005). An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual lexical diversity (MTLD) (Doctoral dissertation). Memphis, TN: University of Memphis <https://www.aaii.org/ocs/index.php/FLAIRS/2010/paper/view/1283>.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392. <https://doi.org/10.3758/brm.42.2.381>
- McCutchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language*, 25, 431–444. [https://doi.org/10.1016/0749-596X\(86\)90036-7](https://doi.org/10.1016/0749-596X(86)90036-7)
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35(1), 13–23. https://doi.org/10.1207/S15326985EP3501_3

- McCutchen, D., Francis, M., & Kerr, S. (1997). Revising for meaning: Effects of knowledge and strategy. *Journal of Educational Psychology*, 89, 667–676. https://doi.org/10.1207/S15326985EP3501_3
- Ministry of Education, Culture, Sports, Science and Technology (2013). English Education Reform Plan corresponding to Globalization. Retrieved from <http://www.mext.go.jp/en/news/topics/detail/1372656.htm>
- Mochizuki, A., Fukasawa, S., Innami, Y., & Koizumi, R. (2015). *Eigo yonginou hyouka no rironto jissen* [Evaluating English 4 skills: Theory and practice]. Taishukan Shoten.
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12, 11–37. <https://doi.org/10.1177/1362168807084492>
- Nation, P. (1991). Fluency and learning. *The English Teacher*, 20, 1–7. https://melta.org.my/journals/TET/downloads/tet20_01_01.pdf
- Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge University Press.
- Nguyen, T. C. L. (2015). Written fluency improvement in a foreign language. *TESOL Journal*, 6(4), 707–730. <https://doi.org/10.1002/tesj.186>
- Nitta, R., & Baba, K. (2014). Task repetition and L2 writing development. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 107–136). Benjamins.
- Norris, J. M. & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. <https://doi.org/10.1093/applin/amp044>
- Obunsha (n.d.). Obunsha Daigakujuken Pasunavi [Obunsha university entrance exam Pasunavi]. Retrieved from <https://passnavi.evidus.com/>
- Ogawa, C. (2016). Examining the effects of types of pre-task planning on oral performances. *JALT Journal*, 38(2), 97–118. <https://doi.org/10.37546/JALTJJ38.2-1>

- Ohi, K. (2016, July 23). *Chuukou eigokyouinno raitinguni kansuru kyoushininchi: Nihon, kankoku, taiwanno eigokyouin taishouhikaku chousakara* [Junior and senior high school English teachers' perceptions on English writing instruction: The comparison of English teachers' surveys in Japan, Korea, and Taiwan]. [Seminar presentation]. JACET SIG on Language Teacher Cognition Colloquium. Tokyo.
- Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing, 19*, 218–233. <https://doi.org/10.1016/j.jslw.2010.10.003>
- Ong, J., & Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers' text quality. *TESOL Quarterly, 47*(2), 375–398. <https://doi.org/10.1002/tesq.55>
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21*, 108–148. <https://doi.org/10.1017/S0272263199001047>
- Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing, 29*, 82–94. <https://doi.org/10.1016/j.jslw.2015.06.008>
- Polio, C. (2012). The acquisition of second language writing. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 319–334). Routledge.
- Polio, C. (2017). Second language writing development: A research agenda. *Language Teaching, 50*(2), 261–275. <https://doi.org/10.1017/S0261444817000015>
- Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge University Press.
- Rostamian, R., Fazilatfar, A. M., & Jabbari, A. A. (2018). The effect of planning time on cognitive processes, monitoring behavior, and quality of L2 writing. *Language Teaching Research, 22*(4), 418–438. <https://doi.org/10.1177/1362168817699239>
- Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing, 28*, 1–19. <https://doi.org/10.1016/j.jslw.2015.02.001>

- Seyyedi, K., Ismail, S., Orang, M., & Nejad, M. (2013). The effect of pre-task planning time on L2 learners' narrative writing performance. *English Language Teaching*, 6(12), 1–10. <https://doi.org/10.5539/elt.v6n12p1>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93–120. <https://doi.org/10.1111/1467-9922.00071>
- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). Cambridge University Press
- Stotsky, S. (1995). The uses and limitations of personal or personalized writing in writing theory, research, and instruction. *Reading Research Quarterly*, 30, 607–632. <https://doi.org/10.2307/748197>
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlehofer (Eds.), *Principle and practice in applied linguistics* (pp. 245–256). 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., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82, 320–337. <https://doi.org/10.1111/j.1540-4781.1998.tb01209.x>
- Tabari, M. A. (2016). The effects of planning time on complexity, accuracy, fluency, and lexical variety in L2 descriptive writing. *Asian-Pacific Journal of Second and Foreign Language Education*, 1(10), 1–15. <https://doi.10.1186/s40862-016-0015-6>
- Tabari, M. A. (2017). Investigating the effects of planning time on the complexity of L2 argumentative writing. *Teaching English as a Second or Foreign Language*, 21(1), 1–24.
- Tanner, P. (2016). Freewriting: Don't think twice, it's all write. *The Language Teacher*, 40(3), 9–12. <https://jalt-publications.org/files/pdf-article/40.3-tlt-art2.pdf>
- Tabachnick, B. G. & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

- Tabari, M. A. (2022). Investigating the interactions between L2 writing processes and products under different task planning time conditions. *Journal of Second Language Writing, 55*, 1. <https://doi.org/10.1016/j.jslw.2022.100871>
- Tavakoli, M., & Rezazadeh, M. (2014). Individual and collaborative planning conditions: Effects on fluency, complexity and accuracy in L2 argumentative writing. *The Journal of Teaching Language Skills, 5*(4), 85–110. <https://doi.org/10.22099/jtls.2014.1857>
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). Second language development in writing: Measures of fluency, accuracy & complexity. In K. Bardovi-Harlig, J. Clark, & J. Puskas (Eds.), *The fluency measures* (pp. 13–32). University of Hawai'i Press.
- Yang, H. C., & Plakans, L. (2012). Second language writers' strategy use and performance on an integrated reading-listening-writing task. *TESOL Quarterly, 46*(1), 80–103. <https://doi.org/10.1002/tesq.6>
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics, 24*, 1–27. <https://doi.org/10.1093/applin/24.1.1>
- Zenker, F., & Kyle, K. (2021). Investigating minimum text lengths for lexical diversity indices. *Assessing Writing, 47*, 1–15. <https://doi.org/10.1016/j.asw.2020.100505>
- Zhang, M. (2022). Collaborative writing in an EFL secondary setting: The role of task complexity. *International Review of Applied Linguistics in Language Teaching*. <https://doi.org/10.1515/iral-2022-0024>

APPENDICES

APPENDIX A
OUTLINING SHEET

Do you agree or disagree?

It is better for university students to live with their parents than to live on their own.

(大学生はひとり暮らしより実家暮らしの方が良い)

Brainstorming & Outlining (5 min.)

Position:

Topic sentence:

1.

Details:

2.

Details:

3.

Details:

Notes-----

APPENDIX B

COMPLETED WORKSHEETS FOR OUTLINING

Do you agree or disagree?

It is better for university students to live with their parents than to live on their own.

(大学生はひとり暮らしより実家暮らしの方が良い)

Brainstorming & Outlining (5 min.)

Position: agree

Topic sentence:

I agree with the opinion.

1. It's not perfectly safe to live alone

Details:

Especially for women, living alone is a little dangerous.

If us live alone, we should tell parents time when we leave home

2. Take cost to life

Details:

pay for the apartment, food etc...

3.

Details:

Brainstorming & Outlining (5 min.)

Position: disagree

Topic sentence:

1. We can spend time freely.

Details:

2. We can learn to live alone.

Details: House work

3. We can easily live alone after I graduate.

Details:

APPENDIX C

PASSAGES FOR THE READING GROUP

Reading 1

Do you agree or disagree? Watching a movie in a theater is better than watching it at home.

(映画は映画館で見る方が家で見るより良い)

5分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I prefer to watch a movie in a theater for two reasons.

First, I can *concentrate on the story if I watch a movie in a theater. If I watch it at home, my family talks to me, or I might get a phone call. So, I must stop it in the middle of the story. I don't like that.

Second, if I watch a movie in a theater, I can feel something special. Watching a movie in a theater in the dark and eating popcorn is very different from everyday life, and I can forget about all problems.

For these reasons, I think it is better to watch a movie in a theater.

Speaker B

I like watching a movie at home for two reasons.

First, I can feel relaxed. I can sit on my favorite sofa in my room or lie on the bed if I want. Sometimes, I don't feel comfortable with the seat at a theater and get very tired. Also, I don't need to care about other *audience at home.

Second, it is cheaper to watch a movie at home. I can watch a movie on Netflix, and it costs only about 1000 yen for one month. If I watch it in a theater, I have to pay 1800 yen for one movie. I can watch many movies on Netflix at home more cheaply.

For these reasons, it is better to watch a movie at home.

*concentrate 集中する

audience 観客

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. 映画は家より映画館で観たい人の方が多い。
2. 話し手 A は、映画館だと突然の来客にじゃまされずに映画に集中できる。
3. 話し手 A は、特別な気分を味わえるから映画館で映画を観たいと思う。
4. 話し手 B は、家で映画を観た方がリラックスできる。
5. 話し手 B によると、Netflix の月額料金は約 1800 円である。

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. More people are willing to watch a movie at a theater than home.
2. Not being interrupted by a sudden visitor, speaker A can concentrate on watching a movie better at a theater.
3. Speaker A wants to watch a movie at a theater because he can enjoy special feeling.
4. For speaker B, it is more relaxing to watch a movie at home than a theater.
5. According to speaker B, the monthly fee of Netflix is 1800 yen.

Reading 2

Do you agree or disagree? Hokkaido is a better place to visit for a vacation than Okinawa.

(休暇で訪れるなら沖縄より北海道が良い)

5 分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I think it is better to visit Hokkaido than Okinawa for a vacation.

First, the weather is comfortable. When I was there last summer, I didn't use the air conditioner at the hotel. I enjoyed the cool weather there in the summer.

Second, Hokkaido is large. Sapporo is a popular city, but it was not crowded. I think that's because it is large. I don't like *the crowd, so I was very happy.

Finally, I can eat expensive sea food *cheaply. I love sushi, especially *Ikura* and *sea urchin. However, I usually can't eat them a lot in Tokyo because they are expensive. In Hokkaido, I can eat fresh sea food less expensively.

In conclusion, Hokkaido is a better place to visit for a vacation.

Speaker B

I prefer to visit Okinawa rather than Hokkaido for the summer.

First, the culture is different from Tokyo. The language is different, and the music is unique. I enjoyed watching an Okinawan dance last time. It was great. Second, the food is very interesting. They eat various kinds of food that is not common in other areas, such as *Umibudou. It's delicious and healthy. Finally, the time is slow in Okinawa. Nobody is in a hurry, and the people are nice and friendly. I can feel very relaxed. It's perfect for a holiday. For these reasons, it is better to visit Okinawa than Hokkaido.

*Umibudou うみぶどう
the crowd 人混み
cheaply 安く
sea urchin うに

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. 話し手 A は、今度の休暇で北海道に行く予定だ。
2. 話し手 A は、動物が好きである。
3. 話し手 A は、刺身が苦手だ。
4. 話し手 B は、沖縄でダンスを鑑賞した。
5. 話し手 B は、沖縄の海で泳いだ。

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. Speaker A is planning to visit Hokkaido for next vacation.
2. Speaker A likes animals.
3. Speaker A does not like raw fish (*sashimi*).
4. Speaker B watched dance performance in Okinawa.
5. Speaker B swam in the sea in Okinawa.

Reading 3

Do you agree or disagree? SNS is a good communication tool.
(SNS はよいコミュニケーションの道具である)

5 分間で以下の文を読み、次のページの問題に答えてください
Read the passage below and answer the questions on next page in 5 minutes

Speaker A
There are a lot of advantages for using SNS.

First, I can keep in touch with my friends easily. My high school friend lives in Hokkaido. Even though we cannot see each other often, I can tell how she is doing by checking her Instagram.

Second, SNS is fun and makes me feel good. When I shared a picture of big pancakes on Instagram the other day, I had 50 “Likes” as soon as I uploaded it. I was very excited.

Finally, it is a good tool to get information. Sometimes, I get information about new restaurants or new shops from my friends’ Instagram.

For these reasons, I think SNS is very useful.

Speaker B

I do not use SNS often for three reasons.

First, it can cause trouble. For example, my classmate shared a photo of me on SNS without my *permission. I was very uncomfortable that my photo was shared with people that I didn’t know.

Second, it is a waste of time. SNS is fun, but I often spend too much time checking and commenting on SNS. I should find better things to do in my life.

Finally, SNS sometimes makes me feel frustrated. The other day, I saw a photo of my friends on SNS enjoying karaoke. I was very *jealous because I was working hard to finish my paper that was due the next morning.

For these reasons, I don’t use SNS.

*permission 許可

jealous 羨ましいと思う

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. インスタグラムは大学生の間で最も人気のある SNS である。
2. 話し手 A は、SNS をお気に入りの俳優の情報を集めるのに使う。
3. 話し手 A は、北海道の写真をインスタグラムに投稿して 50 の「いいね」を得た。
4. 話し手 B は、レポートの締め切りに追われている時に友達がディズニーランドで楽しんでいる写真を見て、羨ましいと感じた。

5. 話し手Bは、自分の誘いを断ったのに別の友達と遊んでいる友達の写真を見てしまい、嫌な気分になった

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. Instagram is the most popular SNS among university students.
 2. Speaker A uses SNS to collect information about favorite actors.
 3. Speaker A posted a photo of Hokkaido and got 50 likes.
 4. Speaker B was jealous when she saw a photo of her friends having a good time at a Disneyland when she was busy working on a paper.
 5. Speaker B felt uncomfortable when she happens to see a photo of her friend playing with another friend inspite of rejecting her offer.
-

Reading 4

Do you agree or disagree? It is a good idea to have pets.

(ペットを飼うのは良いことだ)

5分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

There are a lot of advantages for having pets.

First, pets can *heal people. For example, when I was a child, I *had a fight with my friend at school, but I didn't want to tell that to my mother. When I talked about it to my dog, she *licked my face with a cute smile, and I felt better.

Second, if we have pets, we can make lots of friends. I often post the photos of my dog on Twitter and get lots of *likes from people who love dogs. Then, I started talking about dogs with some of them and *made friends with them.

After a while, we started talking not only about dogs but also about our life. It is fun to have friends who love the same thing.

For these reasons, I think it is good for us to have pets.

Speaker B

I do not think it's a good idea to have pets.

First, it costs a lot to own a pet. If we buy a pet at a shop, it's very expensive. Also, you need even more money to buy food and get *medical treatment if necessary. What would you do if it becomes very sick? We should spend money on ourselves, not pets.

Second, it can cause trouble in the *neighborhood. Dogs are noisy and bite* people. Cats walk through the garden of your neighbors. Some people don't like

animals or have allergy* to them. You need to think about others. You can't be selfish.

For these reasons, I don't agree with having pets.

*heal 癒す

*had a fight ケンカをした

*lick なめる

*like いいね

*made friends with 友達になった

*medical treatment 医療費

*neighborhood 近所

*bite 噛む

*allergy アレルギー

次の文が本文の内容と一致する場合はT、本文の内容と違う場合、あるいは述べられていない場合はFを選択してください

1. 一番人気のあるペットは犬である。
2. 話し手Aは、小さいころ動物が苦手だった。
3. 話し手Aは、Twitterで犬をきっかけに友達を作ることができた。
4. 話し手Bは、ペットより自分自身にお金を使うべきと言っている。
5. 話し手Bは、犬より猫の方が世話が楽だと思っている。

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. Dogs are the most popular pet.
2. Speaker A did not like animals when she was small.
3. Talking about dogs on Twitter, speaker A made friends.
4. Speaker B thinks she should spend more money on herself than a pet.
5. Speaker B thinks cats are easier to take care of than dogs.

Reading 5

Do you agree or disagree? It is good for high school students to have school uniforms.

(高校生にとって制服があるのは良いことだ)

5分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I think school uniform is good. There are three reasons.

First of all, it is good because I do not have to decide what to wear every morning. Personally speaking, I am not interested in fashion and I hate shopping for clothes. So, school uniform makes my life easier.

Second, it saves money because I do not have to buy new clothes every season. I do not get *allowance from my parents. So, I do not have enough money to spend on my clothes. If my school does not have a uniform, I have to get a part-time job to buy nice clothes.

Finally, I like my school uniform because it looks cute. We can only wear it when I am a teenager. It makes me feel special.

For these reasons, I think school uniform is good.

Speaker B

I do not think school uniform is good.

First, I want to wear clothes that I like. I am interested in fashion. So, I want to wear my favorite clothes to school and talk about them with my friends. Also, I want to express myself through fashion.

Second, everyone is different and each person has different *needs. For example, I do not like cold weather, so I don't want to wear a skirt during winter. However, I have to wear it because it is a school uniform. I want to wear *comfortable clothes to school so I can *concentrate on my study.

For these reasons, I do not believe school uniform is good.

*allowance こづかい

*needs ニーズ、必要性

*comfortable 心地よい、快適な

*concentrate 集中する

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. 話し手 A は、ファッションに興味がないし、洋服を買いに行くのも好きではない。
2. 話し手 A は、古着屋をよく利用する。
3. 話し手 A は、制服を着ることで特別な気分になれると感じる。

4. 話し手 B は、友達とファッションの話をしたい。
5. 話し手 B は、勉強に集中するために心地よい服が着たいと考えている

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. Speaker A is not interested in fashion and does not like to go shopping for clothes.
 2. Speaker A often use second-hand clothes shops.
 3. Speaker A feels special when she wears a school uniform.
 4. Speaker B wants to talk about fashion with her friends.
 5. Speaker B wants to wear comfortable clothes to concentrate on studying.
-

Reading 6

Do you agree or disagree? It is better for university students to live with their parents than to live on their own. (大学生はひとり暮らしより実家暮らしの方が良い)

5 分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I think it is good for university students to live with their parents. There are two reasons.

First, their lives will be easier if they live with their parents. Their mothers cook delicious and healthy meals every day, and take care of everything. So, they can focus on their studies or club activities. They can enjoy their campus life.

Second, their families can *save money. If they live on their own, they must spend more money. For example, they have to pay the rent for their apartment, food, and electricity. It costs a lot to live by themselves. Living with their families is cheaper and it makes their parents happier.

For these reasons, it is better for university students to live with their parents.

Speaker B

I think university students should live on their own. There are two reasons.

First, they can learn skills that they need to survive. For example, they can learn how to cook or how to *pay for the bills. It is important to know these skills before they become adults. University students should learn how to be *independent.

Second, they can enjoy freedom. If they live by themselves, they can get up at anytime and come home very late at night. Also, they can invite their friends home *without worrying about their families. It is hard to take care of everything by themselves. However, being free is better than living with their parents.

For these reasons, university students should live by themselves.

*save money お金を節約する

*pay for the bills (公共料金などの)料金の支払い

*independent 自立する

*without worrying about... ...を気にすることなしに

次の文が本文の内容と一致する場合は **T**、本文の内容と違う場合、あるいは述べられていない場合は **F** を選択してください

- 1 話し手 **A** は、一人暮らしだとアルバイトが大変だと述べている。
- 2 話し手 **A** は、実家暮らしのほうが、キャンパスライフを楽しめると述べている。
- 3 話し手 **A** は、一人暮らしに必要な経費として、家賃、食費、電気代の3つをあげている。
- 4 話し手 **B** は、一人暮らしをすることで学べるスキルとして、料理と家計のやりくりをあげている。
- 5 話し手 **B** は、一人暮らしは自由だが危険も伴うと述べている。

(English Translation)

Write **T** if the statement matches with the content of the passage, and write **F** if it is incorrect or not mentioned in the passage.

1. Speaker A says if someone lives alone, one must spend a lot of time for a part-time job.
2. Speaker A says living with parents, one can enjoy university life more.
3. According to speaker A, three necessary expenses for living alone are housing, food and electricity.
4. Speaker B mentions cooking and managing household budget as the skills one can acquire by living alone.
5. Speaker B says one can enjoy freedom by living alone, but it can also be dangerous.

Reading 7

Do you agree or disagree? Paying by electronic money (e.g. pay pay, LINE pay etc.) is better than paying by cash. (電子マネーは現金よりも良い)

5分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I think it is better to use electronic money for three reasons.

First, if I use electronic money, I can get some cash back. For example, if I pay with Pay Pay, I get 1 yen back for spending every 200 yen. I can get a *better deal if I pay by electronic money.

Second, it is easier to pay by electronic money. All I need to do is just to touch the machine with my smartphone. It takes only a few seconds and much quicker than taking out coins in the *wallet. Moreover, coins are actually dirty. It is quicker and cleaner to pay by electronic money.

Third, I don't have to carry a wallet anymore. After I started using electronic money, I can go out only with my smartphone.

For these reasons, I like using electronic money.

Speaker B

I don't think electronic money is good for two reasons.

First, the machines often have problems. For example, I went to a drug store the other day. There was a long line at a cashier because the machine did not scan Pay Pay's barcode. I thought it is much easier to pay by cash.

Second, I do not want to *depend on my smartphone too much. I cannot use electronic money if my smartphone is *out of battery. Also, I might lose my smartphone or *leave it home. Nowadays, I can do anything with my smartphone, but it is also risky. If I do not have a smartphone, I cannot buy anything.

Third, cash is easier to control my *spending. I am afraid I spend too much money if I use electronic money.

For these reasons, I would rather pay by cash.

- *better deal 得な
- *wallet 財布
- *depend on 頼る
- *out of battery 充電切れ
- *leave 忘れる
- *spending 消費

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. 話し手 A は、現金は重くて持ち運びに不便だと述べている。
2. 話し手 A によると、Pay Pay で支払うと 100 円につき 1 円のキャッシュバックがある
3. 話し手 B は、電子マネーは個人情報管理の面で問題があると述べている
4. 話し手 B は、電子マネーが導入されたことで、レジで長いこと待たされた
5. 話し手 B は、現金の方がお金の管理が簡単だと述べている

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. Speaker A thinks cash is heavy and not convenient to carry around.
2. According to speaker A, for every 100 yen spending, one can get 1 yen back for paying by PayPay.
3. Speaker B thinks electronic money is problematic in terms of managing personal information.
4. Having been able to pay by electronic money, speaker B had to wait in the long line at the cashier.
5. Speaker B thinks it is easier to manage one's spending by cash than electronic money.

Reading 8

Do you agree or disagree? Traveling by train is better than traveling by airplane.

(移動手段は飛行機よりも電車が良い)

5 分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I think traveling by train is better than traveling by airplane. There are two reasons.

First, I can enjoy the view from the window. If I travel by airplane, I can only see the sky. However, if I travel by train, I can see mountains, sea, or rice fields. That makes traveling more enjoyable.

Second, a train is more *convenient than an airplane. Usually, an airport is located far from the center of the city. I must take a bus or train to get there. Moreover, I need to get to the airport at least an hour earlier than the departure time to go through a security check. An airplane is faster, but I cannot *actually save much time.

For these reasons, traveling by train is better than traveling by an airplane.

Speaker B

I think traveling by airplane is better than traveling by train. There are three reasons.

First, an airplane is faster than a train. I do not like to travel for long time because it is so *tiring and boring. If I take an airplane, travel time will be shorter. Also, I can play games or watch movies in the airplane. Traveling by airplane makes my trip easier and it is more enjoyable.

Second, traveling by airplane is often cheaper than traveling by train. For example, to go to Hakata, *LCC is cheaper than Shinkan-sen. I can save more money.

Finally, I am collecting mileage points. If I travel by an airplane and collect mileage points, I can get a free flight ticket. I can save a lot of money.

For these reasons, I would rather travel by airplane than train.

*convenient 便利な

*actually 実際

*tiring 疲れる

*efficiently 効率的に

*flight ticket 航空券

*LCC 格安航空会社

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. 話し手 A は、飛行場は都市の中心から離れているので、公共交通機関を使う必要があると述べている
2. 話し手 A は、電車は天候に左右されにくく計画がたてやすいと述べている
3. 話し手 B が、長時間の移動が嫌いなのは、退屈してしまうためである
4. 話し手 B は、博多に行く場合、新幹線の方が LCC よりも安いと述べている
5. 話し手 B は、JAL のマイルポイントをためている

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. Speaker A says one must use public transportation to get to the airport because it is usually far from the center of the city.
2. Speaker A says train is more reliable because it is not easily affected by weather.
3. Speaker B does not like to travel long distance because she gets board.
4. Speaker B says to get to Hakata, Shinkan-sen is cheaper than LCC.
5. Speaker B collects miles for JAL.

Reading 9

Do you agree or disagree? It is good for elementary school children to have smart phones. (giving them access to the Internet) (小学生がスマホを持つのは良いことだ)

5 分間で以下の文を読み、次のページの問題に答えてください

Read the passage below and answer the questions on next page in 5 minutes

Speaker A

I think smartphones are necessary for elementary school children.

First, it is convenient for parents. Smartphones have GPS system. Therefore, parents can check where their children are with smartphones. If their children are late, parents can contact them by making a phone call or sending text messages to check they are safe.

Moreover, according to the survey, more than 50% of the elementary school students have smartphones *nowadays. They communicate with each other using LINE. If children don't have smartphones, they might not be able to join *certain activities with their friends.

Smartphones are necessary for children's safety and for making friends.

Speaker B

Elementary school children do not need smartphones.

First, if they have access to the internet, they might be contacted by dangerous people. They are too young to tell whether someone is a good person or bad person.

Next, it is bad for their health. With smartphones, they have *access to so many interesting games and sites such as YouTube. They are so fun and exciting. Once they start, it is difficult to stop. They stay up late to play games and may not get enough sleep. Moreover, it is bad for their eyesight.

Therefore, elementary school students do not need smartphones.

*nowadays 最近
certain 特定の
access 近づく手段

次の文が本文の内容と一致する場合は T、本文の内容と違う場合、あるいは述べられていない場合は F を選択してください

1. 調査によると今日では約半数以上の小学生がスマホを所有している。
2. 小学生が所有するスマホの多くが「キッズスマホ」と呼ばれるものである。
3. 小学生は LINE でコミュニケーションをとっている。
4. 小学生はスマホで、Instagram などの SNS にアクセスできる。
5. 小学生のスマホはアクセスが制限されているので、事件に巻き込まれる可能性は低い。

(English Translation)

Write T if the statement matches with the content of the passage, and write F if it is incorrect or not mentioned in the passage.

1. According to the survey, more than a half of the elementary school students own smartphones.
2. Many smartphones owned by elementary school students are the models called "Kids smartphone".
3. Elementary school students communicate with each other through LINE.
4. Elementary school students have access to SNS such as Instagram with their smartphones.
5. The type of smartphones owned by elementary school students have limited access to the Internet, so their possibility of being involved in crimes is low.

APPENDIX D

POST-COURSE QUESTIONNAIRE

1. ライティングの前に3通りの事前タスク（アウトラインのみ、口頭でのリハーサル、事前リーディング）を行いました。3つの事前タスクで **timed-writing** を完成させるのに最も役に立ったのはどれだと思いますか。

You have completed timed-writing with three different pre-tasks: outline only, oral rehearsal, and reading. Which task did you find most useful for completing timed-writing?

2. 1.の答えを選んだ理由を書いてください。

Please describe the reason for selecting your response in the question1.

3. 1年間の間 **Timed-writing** を行いましたが、1年間でどんな変化があったか（なかったか）グラフを見ながら簡単に記述してください

You have done timed-writing for one year. Look at your graph and briefly describe how you have changed (or not changed) over one year.

APPENDIX E

INFORMED CONSENT FORM: JAPANESE VERSION

同意書

研究課題名: The effect of pre-tasks on the quality of writing

研究者氏名: 細田 菜穂子

本同意書の目的

本同意書は、研究調査へのご協力をお願いするのためのものです。私は本研究について説明し、ご協力をお願いしたいと思います。研究調査に協力していただくかどうかの決定の前に、質問があれば何でも遠慮なくお尋ねください。この同意書に署名されても、法的権利を放棄することにはなりません。本研究についての説明を受けたこと、質問の機会を得たこと、研究の概要を理解していること、本研究に参加する意思があることを表明することが目的です。

研究の概要

本研究の目的は、異なる事前タスクが、timed-writing の質にどのような影響を与えるか調査することにあります。本研究は私の博士論文の一環として行われますが、将来、規模が大きくなり、出版される可能性があります。研究データは教育および研究目的にのみ使用されます。

協力者へのお願い

本研究に協力していただく期間は 2021 年 春学期、秋学期の 23 週です。この期間中に、timed-writing タスクを 23 回行います。もし参加に同意いただける場合、授業内で書いた作文は研究のデータとして使用されます。

危険性

本研究へのご協力による危険性は最小限だと予想されます。しかしながら、データの利用に関して気が進まない場合はいつでも、お知らせいただければご要望を尊重いたします。途中で参加を辞退しても成績評価に影響を及ぼすことはありません。

研究協力者が得られる利益

本研究へのご協力によって、第2言語によるライティング指導の理解が進み、より効果的な指導法の解明に貢献できます。

研究協力者の権利

協力は任意です。協力を辞退しても構いません。またいつでも無条件で協力の意思を撤回することができます。もし協力を辞退された場合には、あなたに関する情報は本研究に一切記載致しません。本研究に関する質問はいつでもお受けし、納得されるまでご説明できるように最善を尽くします。

守秘義務

参加者の身元及び本研究への協力の守秘に最善を尽くします。本研究により得られた情報は全て安全な場所に保管し、匿名性確保のために最善を尽くします。本研究で使用される人名、施設名、地名などの全ての固有名詞は、個人の特定ができない形で発表されます。参加者の身元が開示されないように最善を尽くしますが、完全な守秘はお約束できません。研究及び参加者の保護に関する規則・規定に私が従っていることを確認するため、大学の施設内倫理委員会などの組織が個人情報を含むデータにアクセスする可能性があります。

データの流布

本研究のデータは私の博士論文に使用されます。また、将来、データは出版物で使用されたり、学会で発表されたりする可能性もあります。

----- (The participants will respond to the statement below via Google form) -----

本研究協力の同意事項

該当するものチェックしてください。

- 本研究への参加に同意しません
- 本研究への参加に同意します

私は、本状記載内容を理解し、本研究に関する私の疑問に対する全ての回答を得ることが出来ました。私は本研究の実施条件を理解し、同意します。

本研究のデータ使用に関する条件がある場合、以下に明記してください。

私は、本研究協力が任意であること、および 細田 菜穂子に伝えることで研究協力をいつでも撤回できることを理解しています。

署名 _____ 日付 _____

ご協力に感謝いたします。

連絡先

住所: AAA

電話番号: BBB

APPENDIX F

INFORMED CONSENT FORM: ENGLISH TRANSLATION

Title of Study *The effect of pre-tasks on the quality of writing*

Researcher *Naoko Hosoda*

Purpose of the Form

The purpose of this form is to ask you to participate in a research study. I will explain the study to you and ask you to volunteer to participate in the study. Please feel free to ask any questions that you may have before you decide if you want to participate. Whether you take part in the study is entirely up to you, and you are also free to withdraw from the study at any time. By signing this consent form, you are not waiving any legal rights that you have. You are only indicating that the study has been explained to you, that you have had an opportunity to ask questions, that you understand what is involved, and that you are willing to participate in the study.

Explanation of the Study

The goal of this study is to investigate how different types of pre-tasks affect the quality of timed-writing. This project is for my dissertation, but it may evolve into a larger study to be published in the future. Data collected in the study will be used for educational and research purposes only.

Request on Participants

The estimated duration of your participation in the study is for two semesters, over 23 weeks. During the period, 23 timed writing tasks will be conducted. If you participate in the study, your compositions will be used as the data for the research.

Risks

The reasonably foreseeable risks or discomforts of participating in this study are minimal. However, if at any time, you do not feel comfortable participating in the study, please let me know and I will respect your requests. The withdrawal from the study will not affect negatively on your grades.

Benefits for the participants

The benefit of participating in this study is knowing that you have contributed to an increased understanding of L2 writing performance and the development of effective teaching method.

Your Rights as a Research Participant

Your participation is voluntary. You may decide not to participate and you are free to stop participating in the research at any time without penalty. If you decide not to participate at any time, information related to you will not be mentioned in this study.

At any time now and in the future, you may ask me questions about this study and I will do my best to answer your questions to your satisfaction.

Confidentiality

I will do my best to keep data collected for this study confidential. I will keep all of this information in a safe place, and do my best to ensure your anonymity. All the proper names (names of people, institutions, places, etc.) used in this study, including yours, will be in a form that cannot be identified. I will also do my best your identity is not to be disclosed, but I cannot promise complete secrecy. In addition, in order to ensure that I am following the rules and regulations regarding research and the protection of human subjects, the IRB may access data including your personal information.

Dissemination

Data collected for this project will go into my Ph.D. dissertation. In the future, the data may also be used for a publication or a presentation at a conference.

------(The participants will respond to the statement below via Google form)-----

Your consent to participate in this study

Please check either box below.

- I do not agree to participate in this study.
- I agree to participate in this study.

I understand the information given to me, and I have received answers to any questions I have about the study. I understand and agree to the conditions of this study.

If you have any other conditions regarding the use of the data provided for this study, please specify them below. _____

I understand that my participation in this study is voluntary, and that I may withdraw from this study at any time by telling Naoko Hosoda to remove me from the study. Contact information is written below.

Name (Please print)

Date

Thank you very much for your participation.

Contact Information

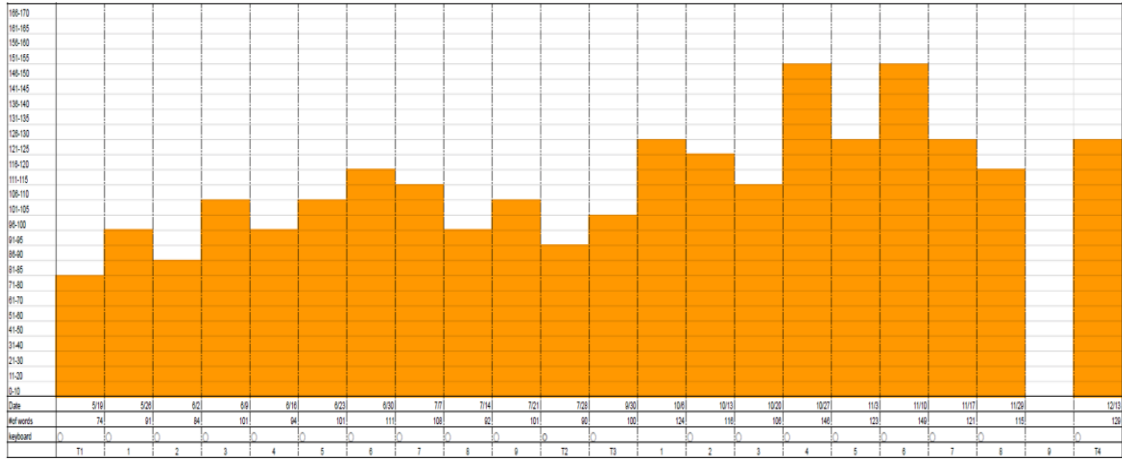
Naoko Hosoda

Address: AAA

Phone: BBB

APPENDIX G

TIMED-WRITING PROGRESS CHART



APPENDIX H

INTERVIEW QUESTIONS

Using the outlining worksheet and composition written just before the interview, please explain what you did and what you were thinking step by step.

- You paused for while here. Why? What were you thinking?
- You deleted some sentences here are replaced with the new idea. Why?
- You wrote this idea on the planning sheet, but did not use it in your writing. Why?
- This idea is not written in your planning sheet. How did you get this idea?

Did you use any ideas or words that are in the reading when you made an outline?
Which one?

Did you use any ideas or expressions that you heard from the partner in your writing?
Which one?

In class, you had opportunities to engage in timed-writing with three pre-task conditions: outlining, oral rehearsal and reading. Which one did you find most useful for timed-writing? Why do you think so?

APPENDIX I

TIMED-WRITING SAMPLE TEXTS

Sample A (Average scores: Content 2.5/ Comprehensibility 2.3)

Yes, I agree. SNS is a good communication tool. Because we can make a lot of friends around the world in order to SNS can update photos and short movies. We have hobbies. It can use communication. SNS can use it. For example, we use a chat called 'Direct Message'. It can chat with a lot of people.

Sample B (Average scores: Content: 3.5/ Comprehensibility: 4.0)

I think it is better for university students to live with their parents than to live on their own. In fact, I live on their own. I have three reasons that it is better for university students to live with their parents than to live on their own. Firstly, living costs are much less than living on your own. We will have to pay rent, water and electricity charges once a month. In fact, I pay as much as 50000 yen once a week. Secondly, we may be sad alone. I speak to no one at home. But I am not very sad about my parents and brother. For these reasons, I think it is better for university students to live with their parents than to live on their own.

Sample C (Average scores: Content: 5.0/ Comprehensibility: 5.5)

I agree with this opinion that it is better for university students to live with their parents than to live on their own. I have three reasons why I think so. First of all, it doesn't cost money. For example, they don't have to pay for food, water, electricity, and so on. The second is that they are so busy that it is difficult to do housework because they don't have much time to study or work part-time. And cleaning and washing in a short amount of time is difficult. Third, there will be someone at home every day. If they live alone, there is no one at home but If you live in your parents' house, there should be someone at home. So, they are less likely to be homesick and they are not lonely. That is why I think that it is better for university students to live with their parents than to live on their own.