



Effects of Pre-Task Planning and Source Texts on Korean EFL College Learners' Summary Writing

Yoonhee Choe

Chongshin University, South Korea

Jisu Ryu

Konkuk University, South Korea

Moongee Jeon

Konkuk University, South Korea

This study aims to investigate to what extent pre-task planning and the source text type(genre) affect Korean English as a Foreign Language (EFL) college learners' summary writings in terms of lexical, sentential, and discourse-level features. A total of 120 summary writings of cause/effect expository texts and argumentative texts in the different modes of planning were collected and analyzed using a computational assessment tool, Coh-Metrix. The results show that the participants' summary writings in the different planning conditions and text types were statistically different according to their lexical-level (the mean word length, word frequency, imageability, concreteness, the third person pronouns), sentential-level (the mean sentence length, causal connectives, temporal connectives, noun density, Flesch-Reading Ease (FRE), Flesch-Kincaid Grade Level (FKGL)), and discourse-level (type-token ratio, LSA cosines for all and adjacent sentences) features. This study provides some pedagogical implications for teaching English summary writings of various planning contexts and source text types.

Keywords: pre-task planning, no planning, source text types, summary writing, Coh-Metrix

Introduction

Writing is a complex linguistic ability that requires more cognitive processing as a productive skill than other receptive abilities like reading and listening in a second or foreign language (Li, 2014; Yoon & Polio, 2017). Writing encompasses not just independent writing, which consists of merely writing facts or thoughts on a particular topic, but also integrated writing, which involves reading or listening to a source text and then summarizing it (Johnson et al., 2012; Li, 2014; Yoon & Polio, 2017). The importance of integrated writing, including authentic activities such as source reading, reorganizing, and summarizing, is increasing in school classrooms and in the field of English assessment. When compared to independent writing activities, integrated writing tasks such as *discourse synthesis*, *reading to write*, *sourcing*, and *summary writing* are perceived as more challenging (Kirkpatrick & Klein, 2009, p. 309).

Since these integrated reading and writing activities in English are known for their authenticity, they have been incorporated in several standardized official English proficiency tests such as the IBT TOEFL



(e.g., Choe, 2013; Li, 2014; Plakans, 2009; Yang & Plakans, 2012; Yu, 2009). In addition, several comparison studies have been carried out to examine how the pattern of integrated writing changes based on the type (i.e., genre) of source text read by learners. Learners' summary writings, as an example of integrated writing, differ depending on the genre (i.e., type) of writing, according to previous studies (Choe et al., 2020; Jeon & Choe, 2019; Li, 2014). Jeon and Choe (2019), for example, analyzed the linguistic features of college students' summary writing manipulating the genre of source texts (argumentative vs. expository texts) with Coh-Metrix, a language analysis tool. In their argumentative summaries, the learners used more words, low-frequency and concrete terms, adjectives, and longer sentences. In their explanatory summaries, however, the participants used more nouns and verbs. The learners also used more conjunctions and syntactically complex sentences in their argumentative writing. The referential coherence score for adjacent sentences was also higher in the argumentative summaries. The results of Choe and Jeon's study suggest that the genre type of the source text affects students' summary writing differently, indicating that the linguistic characteristics by the software are reflected in their summary writings in a different manner depending on the genre of the source texts. In addition, Choe et al.'s follow-up research (2020) examined how the three types of source text (i.e., argumentative texts vs. cause/effect expository texts vs. comparison/contrast expository texts) influenced participants' summary compositions. Their findings showed that there were significant differences among the three text types for the lexical level (mean word length, word frequency, imageability, concreteness, and third person pronouns), the sentential level (mean sentence length, causal connectives, temporal connectives, noun density, FRE, and FKGL), and the discourse level (type-token ratio, LSA cosines for all) measures. Choe et al.'s study additionally showed that the original source texts also influenced the summaries written by the students. Other studies also showed that students' writing habits differed based on the source text's vocabulary level, the complexity of the sentence structure, and the discourse structure. (e.g., Li, 2014; Yu, 2009). These studies suggest that the type of source texts has different effects on students' summary writings. These studies also illustrate how these integrated writing activities are more cognitively difficult and sophisticated than independent writing activities (i.e., simple standalone writing tasks).

Another important aspect of L2 writing that draws researchers' attention is the influence of pre-task planning. Much research has looked at the effects of pre-task planning on L2 performance as an important factor of L2 task designs over the last few decades (Fazilatfar et al., 2020; Johnson et al., 2012; Kang & Lee, 2019; Moon, 2019; Tabari, 2020). The effect on planning has been shown to be highly consistent in L2 oral output (e.g., Yuan, & Ellis, 2003). Numerous studies on the influence of pre-task planning in L2 independent writing have also been conducted for investigating its effects on L2 performance, but the consequences of previous studies on planning in L2 writing have not been always consistent (e.g., Johnson et al., 2012; Kang & Lee, 2019).

Besides, research on planning in integrated writing has just begun to emerge, and there is still much more work to be done for systematically revealing the effect of planning on L2 performance. Previous research has produced noteworthy achievements by examining the effects of genre and source on students' performance in different writing learning modes (i.e., independent writing vs. integrated writing). However, these studies mainly focus on investigating independent writing activities to reveal the effects of genre and source on L2 learners' performance. This has prompted several researchers (e.g., Choe et al., 2020; Jeon & Choe, 2019; Li, 2004; Yu, 2009) to examine the effects of genre and source on L2 performance in integrated writing. However, little attention has been paid to the influence of planning in integrated writing tasks, compared to that in independent writing tasks (Payant et al., 2019).

According to previous research, pre-task planning has varying effects on independent writing in terms of accuracy, complexity, and fluency. Integrated writing demands the development of progressively sophisticated cognitive skills for comprehending the original text as well as paraphrasing and summarizing it. Thus, it is necessary to do research for exploring the extent to which pre-task planning influences integrated writing because pre-task planning may have a greater impact on complicated tasks than on easier tasks (Ellis, 2022; Kang & Lee, 2019).

Driven by the implications and limitations of the previous studies, the current study aims to focus more on investigating the main effect of pre-task planning and the interaction effect of planning and genre (i.e., source text types), particularly in the integrated writing environment. This study also attempted to investigate the effect of sources on planning across different genres. One thing to note is that the current study, which aims to analyze the effects of planning and sources on L2 performance (i.e., writing ability) in two different genres in integrated writing, can be differentiated from previous studies that only consider a single genre such as narration or argumentation.

Considering the results of previous studies and the main purpose of this study, the research questions of the present study are as follows:

1. What effect does pre-task planning have on the summary writing of L2 learners in the integrated writing environment, according to Coh-Metrix analysis?
2. What effect (i.e., interaction effect) do pre-task planning and genres (types of source text) have on the summary writing of L2 learners in the integrated writing environment, according to Coh-Metrix analysis?

In addition to these major research questions, this study presents the results of analyzing the effects of source texts on pre-task planning across two different genres, based on descriptive statistics.

Literature Review

Effects of Planning in L2 Writing

In L2 writing, pre-task planning refers to the amount of time allotted to contemplate the topic of the essay or to create an overarching strategy prior to writing (Johnson, et al., 2012; Ortega, 1999). Based on information processing theory, pre-task preparation may assist L2 writers in concentrating on translating a source language into a target language (i.e., turning one's thoughts into words through the use of target language lexical items and grammatical rules) throughout the writing process, particularly for novices (Oh et al., 2015). This can be explained by the concept of restricted working memory, which is traditionally defined as the temporary storage (i.e., short-term memory) and has come to refer to a cognitive process necessary to actively manipulate the stored information (Baddeley, 2003; Johnson, et al., 2012; Kellogg, 1996; Moon, 2019). L2 learners generally write engaging in complex and demanding cognitive processes such as generating and selecting ideas, defining goals and organization, and following out those plans in L2 using suitable syntactic structures and terminology in their writings (Johnson et al., 2012; Oh et al., 2015; Ong & Zhang, 2010). L2 learners cannot give equal attention to the various cognitive processes required for writing due to the limited capacity of working memory (Kellogg, 1996). According to Ellis and Yuan (2004), pre-task planning time reduces the demands on working memory resources, allowing L2 writers to focus more on translation (i.e., from a source to a target language) and other cognitive writing processes. If L2 writers are required to write without planning, they must plan, translate, execute, and monitor all at the same time (Johnson et al., 2012). Hence, it is an effective writing strategy to give L2 learners planning time so that they can efficiently allocate the limited resource of working memory to ideas and organization, contributing to producing well-formed writing outcomes throughout the writing process (Ellis & Yuan, 2004).

Many scholars have tried to reduce the cognitive information processing load on second language learners (e.g., Johnson, 2017). Among other characteristics, the measures for CAF (complexity, accuracy, fluency) and lexical sophistication were statistically examined to see whether there were differences in communicative competence between pre-task preparation, on-line planning tasks versus spontaneous communication tasks while completing them. Ellis and Yuan (2004), in their L2 writing study, looked at how planning affected the CAF measures for writings produced by Chinese EFL university students.

Planning enhanced the complexity and fluency of language of the L2 learners, according to their findings. On the other hand, planning had no influence on the linguistic accuracy. Such influence of planning on fluency was also verified by Johnson et al. (2012). A more recent study, however, found that this result was not always convincing; planning, according to Ong and Zhang (2013), can impair fluency and lexical complexity in the writing of L2 learners.

Several research on the effects of pre-task planning on L2 writing has been undertaken in Korea. Oh et al. (2015) examined the relative contributions of planning, L2 linguistic competence, and individual differences to the argumentative essays of Korean EFL students. They conducted a series of hierarchical multiple regression analysis. According to their findings, planning alone lacked the predictive power necessary to adequately explain the quality of L2 writing. They suggested that a substantial contribution to L2 writing was discovered to be participants' capacity to recognize effective L2 essays and the content quality of the outline created during planning. Grammar, productive vocabulary, and sentence processing speed in L2 linguistic expertise, as well as individual differences such as self-efficacy, strategy, and motivation, all contributed more significantly to the participants' L2 writing. More recently, Kang and Lee (2019) investigated the impact of individual versus collaborative pre-task planning on L2 writing. Collaborative planning, they discovered, is more effective in terms of syntactic complexity and fluency but not accuracy. However, in more complicated picture description tasks, complex and individual planning were comparable in terms of lexical complexity, whereas individual planning was more effective in simpler tasks.

Although previous studies have been conducted to investigate the effect of planning on L2 learners' writing, their results are not always consistent, presumably due to the differences in the target language of the participants (e.g., Spanish, and Chinese) and in their knowledge of the writing types (i.e., genre). Other task variables, such as the complexity of the writing task (simple vs. complex), the planning structure (structured vs. unstructured), and the planning style (individual vs. collaborative), also had a significant effect on the planning effects on L2 writing.

However, there are previous studies on planning that show somewhat consistent results that complexity and fluency are improved over accuracy. Tabari (2016), for example, investigated the lexical diversity of L2 learners' writings in addition to the CAF measures (i.e., complexity, accuracy, fluency) mainly examined in previous studies. His research revealed that L2 learners who completed their writing under the pre-task planning condition created more fluent texts than those who did not prepare their writings; the L2 learners with no pre-task planning also tended to generate texts with a greater lexical variety. His research postulated that there was a trade-off between fluency and lexical variety.

Considering the implications of previous studies, the current study particularly focused on analyzing several measures of Coh-Metrix related to the CAF and lexical diversity features examined by previous studies to reveal the effect of planning on L2 learners' summary writings.

As indicated previously, academia has recently begun to focus more on the effect of preparation on summary writings, an integrated writing activity utilized on the TOEFL and CAEL tests. Payant et al. (2019) conducted a unique study to assess the effects of source comprehension, individual differences, and prewriting planning on the performance of English L2 writers on an integrated writing assignment administered by the CAEL Assessment. The strategies and the degree of elaboration used in participants' prewriting planning notes were studied. Their study showed that the influence of the source (i.e., source comprehension) was extremely significant for the L2 learners' writing performance, suggesting that the source effect should also be investigated further in order to comprehensively determine the effect of planning on the integrated task. They argue that it is critical to perform a more thorough examination of the influence of planning together with source materials, when undertaking an integrated summary writing.

As suggested in the preceding studies, pre-task planning in L2 writing has a significant impact on learners' writings linguistically. Additionally, there is insufficient research that has attempted to examine the linguistic features presented in summary writings at the word, sentence, and discourse levels across given source text types/genres. Accordingly, the current study, unlike previous studies, applied Coh-Metrix, a multilevel language analysis tool, to explain the effect of planning on Korean EFL students' summary writings extensively, according to genre and source.

Interaction between Writing Genre and Planning Time

Researchers have examined the effects of planning on writing performances in different genres and yielded notable findings (Ong & Zhang, 2010; Rahimi & Zhang, 2018; Tabari, 2017 for the argumentative genre; Kirkpatrick & Klien, 2009 for the compare/contrast genre; Seyyedi et al., 2013; Rahimpour & Safarie, 2011; Tabari, 2016 for the narrative/descriptive genre). For example, Tabari (2017) investigated the effects of different planning conditions (pre-task, online, no planning) on the syntactic complexity, diversity, and lexical variety of the 90 EFL undergraduate students' argumentative writings. Those in the three planning conditions were provided with the same amount of time for writing (i.e., 17 minutes). However, the time allotted for planning before the task was different across the conditions; that is, the participants in both no planning and on-line planning conditions were provided with 0.5 minute to plan beforehand, whereas those in the pre-task planning condition were given 10 minutes. The main results are that pre-task planning has a significant effect on syntactic complexity (measured by the proportion of clauses to T-units) and syntactic variety (calculated through the number of different grammatical verb forms) while the online planning has a positive impact on syntactic complexity. Notably, both planning groups did not benefit from additional planning time in terms of the lexical variety (as indicated by the type-token ratio).

Similar findings were obtained from Rahimi and Zhang (2018) in which the effects of task complexity and planning conditions were analyzed on EFL learners' argumentative writings. They had 80 EFL learners of upper-intermediate proficiency complete two writing tasks of varying degrees of complexity under pre-task versus no pre-task planning conditions. In line with the findings reported in Tabari (2017), pre-task planning was shown to enhance syntactic complexity (measured by the number of subordinate clauses per clause) but not to have a positive influence on the lexical diversity and academic word use. Previous studies in planning time in the argumentative genre have produced some consistent, as well as mixed, effects in terms of CAF (i.e., complexity, accuracy, and fluency). Overall, mixed effects are observed on fluency, whereas positive effects on syntactic complexity, and no significant effects on accuracy and lexical diversity (or complexity) are consistently found.

Other researchers are more concerned with the effects of planning time on L2 narrative/descriptive writing performances. For example, Seyyedi et al. (2013) examined the effects of pre-task planning time on Malaysian EFL students' narrative description while looking at a set of pictures. The main findings showed that pre-task planning led to greater fluency (measured through the number of words per minute) and lexical density (measured through the ratio of lexical words to total words), but no substantial effects were discovered on the accuracy (i.e., the number of the error-free clauses). In a similar study, Tabari (2016), in which 78 intermediate EFL learners in three planning groups (i.e., pre-task planning, online planning, and no planning) completed a picture description task, also revealed that pre-task planning promoted higher fluency (estimated via the number of syllables), whereas no significant effects are identified on the accuracy (measured through the error-free clauses). However, pre-task planning was not effective for enhancing lexical variety (estimated by the type-token ratio) and complexity (the average T-unit length), which is inconsistent with the previous finding (i.e., Seyyedi et al., 2013). Taken together, prior studies have demonstrated that L2 learners, provided with pre-task planning, produce narrative compositions of greater fluency while pre-task planning does not have a substantial effect on accuracy, lexical variety, and complexity.

Method

Participants and Procedures

A total of 60 students in grades 3 and 4 who were majoring in English education at a university in Seoul participated in this study. In terms of the CEFR framework, their band of English proficiency ranged from B2 to C1. Thirty of them were in the control group (i.e., no-planning group), and each had to write two summaries for argumentative and expository texts (i.e., source texts) in 30 minutes without any prior planning. Before summarizing, as an experimental group, another 30 participants were given 20-minute planning and preparation time in advance. Previous research has revealed a wide range of pre-task planning time intervals. Occasionally, ten minutes are allotted for pre-task planning and seventeen to thirty minutes for task completion (e.g., Ellis & Yuan, 2004; Johnson et al., 2012). The earlier studies relied heavily on independent writing, which was quite simple to accomplish. However, in this study, participants are required to perform an integrated task that requires them to read the original text first and then summarize it. That is why the experiment group was given a time limit of twenty minutes to complete the activity.

Prior to the study, the writing tasks assigned to participants in each group were gathered and graded to confirm that the two groups were equal. Between the two groups, there were no significant differences in writing scores. Both groups had a similar degree of genre knowledge and the participants in two groups gained a basic understanding of each genre prior to the test.

The participants completed their tasks in class using their pencils. Before summarizing, the experimental group was expected to plan in terms of content, organization, and language. In their sheets, they were given a pre-task planning box in which they were instructed to create comments in their own words and outline the structure. The control group was not required to perform any of these pre-task planning activities. They just did their tasks while reading and summarizing the source task. To minimize the order effect in this study, half of the participants in each group completed the argumentative task first and the other half completed the cause-and-effect task first.

The expository source text was about “Causes and Effects of Melatonin on Human Body”, while the argumentative source text was about “Can Human Teachers be Replaced by A.I. Teachers?” The source texts were adapted for this study from an English teachers’ placement exam in Korea.

Data Collection and Analysis

All learners’ written compositions were saved as separate txt files to build corpora. A total number of text files is 120, each group comprising of 30 files. These txt files are then subject to linguistic analysis. The specific linguistic measures of the software are provided in the next section. For the statistical analyses, two-way between-subjects analysis of variance (ANOVA) techniques were performed. The independent variables are *planning* (planning vs. no planning) and *genre* (argumentative vs. cause-and-effect expository text), whereas the dependent variables are each of the Coh-Metrix measures. The significance level was set at 0.05 ($p = .05$).

Coh-Metrix measures

Word-level measures. The word-level measures by Coh-Metrix include the mean word length, word frequency, lexical diversity, word feature, and pronoun values. The mean word length indicates the number of characters divided by the number of syllables; longer words tend to increase the difficulty of text (Grasser et al., 2004). The word frequency measure indicates a logarithmic-transformed word frequency value based on the CELEX lexical database made by the Dutch Centre for Lexical Information (Baayen et al., 1995). The logarithmic-transformed word frequency measure tends to be more appropriate when applied to text analysis (Graesser et al., 2004). Specifically, the distribution of frequency scores is

not close to the normal distribution. Namely, the original frequency scores by themselves are not suitable for statistical analysis based on a linear model. However, since the distribution of logarithmic-transformed frequency scores tends to approximate the normal distribution, the logarithmic-transformed frequency scores are more suitable for statistical analysis (Graesser et al., 2004).

The lexical diversity measure refers to a type-token ratio. The type value is the frequency of a word first presented in the text, so the type value of the word always should be 1 (Graesser et al., 2011). The token value indicates the frequency of a word repeatedly used in the text. The higher this value is, the more difficult it is to understand the text.

The word feature measures consist of imageability, concreteness, and age of acquisition values. The Coh-Metrix system applies the MRC Psycholinguistic database for computing these measures (Coltheart, 1981); these values are presented in the range of 100 to 700. In the case of imageability and concreteness measures, the higher these values for a word are, the lower the understanding of the word is, whereas if the age of acquisition value of a word is high, it is relatively difficult to understand the word because the word is learned at a higher age (Graesser et al., 2011).

Pronoun is an important linguistic factor that influences text comprehension. It can contribute to building text cohesion, but a text with too many pronouns can be difficult to understand because text readers have to find the referents to which these pronouns refer in order to resolve their meanings (Graesser et al., 2004). As such, pronouns have an important influence on text processing, so they can affect the summary writings of the current study, according to genre and planning. The first-person pronouns, in particular, can be used more in the planning condition where the participants of the planning condition can fully express their thoughts self-centeredly across genres. Coh-Metrix provides the first-person, the second-person, the third-person, and all personal pronoun measures. They are the relative frequency values of personal pronouns appearing per 1,000 in the text (Graesser et al., 2004).

Sentence-level measures. The sentence-level measures consist of the mean sentence length, standard readability indices, syntactic complexity, and the relative frequency scores for connectives. The mean sentence length indicates the number of words divided by the number of sentences in the text. Longer sentences tend to make the text difficult to understand (Graesser et al., 2004).

The standard readability indices are the Flesch Reading Ease (FRE) score and the Flesch-Kincaid Grade Level (FKGL). The word length and the sentence length determine each measure (Graesser et al., 2004). The FRE score ranges from 0 to 100; the higher this value of a text, the easier it is to understand the text. On the other hand, the FKGL score corresponds to the US grade level ranging from 1 to 12. A text with a higher FKGL value is difficult to understand.

Syntactic complexity includes the number of words before main verbs and the value of noun phrase (NP) density. The NP density is the ratio of modifiers to all words in the NP (Graesser et al., 2004). A syntactically complex text is difficult to understand. So, the higher these values of a text are, the more difficult it is for readers to understand the text (Graesser et al., 2011).

The connective measures by Coh-Metrix have causal (e.g., so, because, etc.), additive (e.g., but, moreover, etc.), temporal (e.g., until, first, etc.) connectives. These are the relative frequency scores of connectives occurring per 1,000 words. Since connectives, especially causal connectives, can facilitate the connections between sentences, they play an important role in forming text cohesion (Halliday & Hasan, 1976).

Discourse-level measures. The discourse-level indices by Coh-Metrix include co-referential and semantic cohesion scores (Graesser et al., 2004). The co-referential cohesion measure is the argument (nouns, pronouns, NPs) overlap for adjacent sentences or for all sentences. The semantic cohesion measures consist of a LSA (Latent Semantic Analysis) score for adjacent and all sentences. LSA is a statistical model applied for measuring the semantic similarity between two text units based on a large corpus data. These cohesion measures have a significant effect on text comprehension. If the cohesion of a text is higher, it is easier to understand the text (Landauer et al., 1998).

Results and Discussion

Preliminary Analyses

Previous studies (e.g., Choe et al., 2020) have suggested that the presence of the source text has a significant influence on the quality of the summary writings. To gauge the overall impact of the source text on the summary writings, this study conducted the preliminary concordance-based keyword analyses before the main analyses. The procedure for the preliminary analyses is as follows. First, the keywords (specifically, key *content* words), used more than once, were identified in the source text from most to least in order of frequency. Then, for the summary writings in the planning and no planning conditions, a frequency analysis was made of the list of the words extracted from the keyword analysis on the source text. This was conducted for argumentative and cause and effect expository essays separately. The results are presented in Table 1 (for argumentative essays) and Table 2 (for cause-and-effect essays).

As shown in Table 1, the keyword analysis on the source text in the argumentative genre reveals that “teachers” are used most often (11 times) in the source text, followed by “students” (6), “artificial” (4) and “intelligence” (4), “human” (3), “need” (3), and “skills” (3). Other words (e.g., classroom, digital, experts) are used twice. A frequency count of the keywords, extracted from the source text, was made of the texts in planning and no planning conditions. The results showed that all of the keywords in the source text were used more than 10 times in the writings produced in the no planning condition (on the rightmost column), except for “experts,” “schools,” and “worry”, whereas only three words, “teachers,” “students,” and “intelligence” were adopted in the writings produced in the planning condition (on the second column). In fact, over half of the keywords in the source text were not even used in the planned writings. Finally, note that in both conditions, the most frequently used word in the source text, “teachers” (11), were used most often (54 and 112 respectively). The current keyword analysis reveals that the source text influences the writing performances, particularly when the students are not provided with planning time beforehand. This is presumably due to the lack of planning time and external resources, such as the internet or the dictionary, thereby leading to the overreliance of the students on the source text and the keywords directly derivable from the text.

TABLE 1
The Effect of Source on Planning Time for Argumentative Essays

Content words	Planning (n = 30)	Source (n = 1)	No planning (n = 30)
teachers	54	11	112
students	14	6	96
artificial	0	4	58
intelligence	25	4	57
human	0	3	95
need	3	3	24
skills	1	3	44
classroom	0	2	18
digital	0	2	15
experts	1	2	9
help	2	2	16
learning	4	2	12
obsolete	0	2	10
resources	0	2	13
robots	0	2	30
schools	0	2	8
worry	0	2	4
No. of Tokens	104	54	621

Table 2 presents the results of the analysis on the texts in the cause-and-effect expository genre. The keyword analysis on the source text yielded 16 key content words which were used more than once. The most frequent word in the source text “melatonin” were also most frequently used in the texts in both

planning (77 times) and no planning conditions (152 times), which is consistent with the previous analysis for the argumentative writings. In addition, all of the identified keywords were adopted by those in the no planning condition whereas those in the planning condition adopted only some of the keywords. Also, the total number of tokens indicates that the students in the no planning condition used twice as many keywords as those in the planning condition, again reflecting their heavy reliance on the source text.

The descriptive comparison between the two genres demonstrated that the learners in the planning group used more of the keywords in the cause-and-effect genre than in the argumentative genre. The inspection of the keywords between the two source texts suggests that the keywords in the cause-and-effect genre include more technical terms (e.g., “melatonin”, “proteins”, “Alzheimer”, and “hormone”) which are therefore more difficult to replace with other synonyms. All in all, the interaction of the source text and the planning variable is more prominent in the argumentative than the cause-and-effect genre. That is, the differences between the two planning conditions are more pronounced in a certain genre (i.e., argumentative).

TABLE 2

The Effect of Source on Planning Time for Cause & Effect Essays

Content words	Planning (n = 30)	Source (n = 1)	No planning (n = 30)
melatonin	77	7	152
cognitive	28	3	43
cycle	11	3	33
disorder	10	3	32
proteins	7	3	20
research	0	3	15
toxic	5	3	24
affective	9	2	22
alzheimer	11	2	22
effects	14	2	22
hormone	10	2	28
mismatch	0	2	10
order	0	2	33
patients	2	2	7
people	14	2	35
seasonal	13	2	23
No. of Tokens	212	45	536

Main Analyses

A series of analysis of variance (ANOVA) techniques were performed, with planning and genre as between-subjects independent variables and 22 Coh-Matrix measures as dependent variables. It should be noted that the current study is to examine the effects of planning and the interaction effects of planning and genre because a previous study has revealed the main effect of genre (Jeon & Choe, 2019). So, this study did not consider analyzing the main effect of genre.

Word-level analyses

Table 3 presents the analyses on the summary writings in two genres (cause and effect vs. argumentative text) in two planning conditions (planning vs. no planning) at the word-level.

According to the two-way ANOVA, there are no significant main effects of planning and interaction effects of planning and genre on word length, age of acquisition, second-person pronouns and third-person pronouns (all $ps > .05$). However, there are main effects of planning on type-token ratio ($F(3, 116) = 30.470, p < .01$) and first-person pronouns ($F(3, 116) = 12.455, p < .01$). The examination of the mean scores reveals type-token ratio was greater and therefore lexical diversity is higher in the planning condition. On the other hand, first-person pronouns are more used in the no planning condition. The conflicting findings with the previous ones (e.g., Ong & Zhang, 2013; Tabari, 2016) can be accounted for

by the presence of the source text in the current study. When writing an argumentative essay without a source text (i.e., independent writing) as in Tabari's research (2016), learners' lexical repertoire may not significantly vary with additional time. However, when learners are provided with a source text, the words used in the source text serve as a retrieval cue to other related words (e.g., synonyms). Thus, the learners of the planning condition may have taken advantage of the additional planning time to have other relevant words (i.e., synonyms) activated. The present study implies that the effects of planning can differ by the type of tasks (i.e., independent writing vs. integrated writing).

For word frequency, imageability and concreteness, interaction effects of planning and genre, as well as main effects of planning, were observed. Due to the interaction effects, simple main effects were analyzed. All three lexical measures did not display significant differences between the two planning conditions in the cause-and-effect genre, whereas all three yielded substantial differences in the argumentative genre (all $ps < .01$). Specifically, those in the planning condition included words of higher frequency but lower imageability and concreteness than those in the no planning condition.

TABLE 3

Word-level Analyses on L2 Summary Writings as a Function of Planning and Genre

	Cause and effect (n = 60)			Argumentative (n = 60)		
	Planning (n = 30)	No planning (n = 30)	Total (n = 60)	Planning (n = 30)	No planning (n = 30)	Total (n = 60)
Basic Counts						
mean word length	1.72(.11)	1.72(.13)	1.72(.12)	1.62(.12)	1.70(.12)	1.72(.12)
Lexical Features						
word frequency	2.25(.20)	2.26(.20)	2.26(.20)	2.57(.12)	2.38(.14)	2.47(.16)
type-token ratio	.88(.06)	.83(.10)	.86(.09)	.82(.07)	.71(.09)	.76(.10)
imageability	386(22.9)	382(25.9)	384(24.3)	360(12.3)	403(18.8)	381(26.9)
concreteness	356(22.2)	355(25.8)	355(23.9)	325(14.9)	375(19.2)	350(30.5)
AOA	374(32.4)	363(28.9)	359(30.9)	375(29.7)	367(24.3)	371(27.2)
Personal Pronouns						
1 st person pronouns	1.34(3.55)	5.29(8.51)	3.32(6.76)	.76(2.44)	3.60(4.50)	2.18(3.86)
2 nd person pronouns	1.64(6.39)	2.55(9.64)	2.09(8.12)	1.28(4.58)	.57(2.00)	.92(3.52)
3 rd person pronouns	.17(.93)	1.85(4.09)	1.01(3.06)	12.1(6.60)	11.5(8.09)	11.8(7.33)

The current findings that the learners' writing performances are not significantly different by planning time in the cause-and-effect genre are consistent with the results of the preliminary analyses. The preliminary analyses showed that the learners relied more heavily on the source text in the cause-and-effect genre than the argumentative genre, regardless of the planning time. This is largely due to a high proportion of the technical terms included in the cause-and-effect expository genre, leading to the excessive reliance on the source text in both planning conditions. This accounts for the negligible effects of the planning time on the learners' performances in the cause-and-effect genre. On the other hand, those in no planning condition appeared to copy many words from the source text which includes words of low frequency and higher imageability and concreteness. However, those in the planning condition were given additional planning time to come up with more difficult-to-imagine and abstract words (i.e., lower imageability and concreteness), albeit higher in frequency. The present findings on the word-level analyses expand the previous research on planning by demonstrating how planning time can also affect lexical features such as frequency, imageability, and concreteness.

Sentence-level analyses

Sentence-level analyses include analyses of the mean sentence length, FRE, FKGL, NP density, subject density, causal connectives, additive connectives, temporal connectives, and all connectives. The results on the sentence-level analyses are shown in Table 4. The analyses show that there are no main effects of planning as well as interaction effects on sentence length, temporal connectives, and all connectives (all

$ps > .05$). The main effects of planning, however, are found in measures of subject density ($F(3, 116) = 5.826, p = .017$) and causal connectives ($F(3, 116) = 4.549, p = .035$). The planning group produced more complex structures including longer subjects than the no planning group. The no planning group, on the other hand, used more causal connectives than the other group.

Interaction effects are obtained on FRE ($F(3, 116) = 5.537, p = .020$) scores, FKGL scores ($F(3, 116) = 5.624, p = .019$), NP density scores ($F(3, 116) = 10.290, p = .002$), and additive connectives ($F(3, 116) = 10.976, p = .001$). Given the interaction effects, simple effects were examined instead of the main effects of planning. All of the four sentence-level measures were shown to be significantly different between the two planning conditions only in the argumentative genre (all $ps < .05$), whereas not in the cause-and-effect genre (all $ps > .05$). When provided with planning time, the learners produced more complex writings than those without planning time, as evidenced in the lower FRE scores, higher FKGL scores, and greater NP density scores. On the contrary, those in the no planning condition made a more intense use of the additive connectives.

Table 4

Sentence-level Analyses on L2 Summary Writings as a Function of Planning and Genre

	Cause and effect (n = 60)			Argumentative (n = 60)		
	Planning (n = 30)	No planning (n = 30)	Total (n = 30)	Planning (n = 30)	No planning (n = 30)	Total (n = 60)
Basic Counts						
mean sentence length	18.1(5.01)	18.5(4.34)	17.1(4.37)	18.6(4.34)	18.5(4.34)	18.6(4.30)
Standard Readability						
FRE	42.7(9.62)	44.9(10.4)	43.8(10.00)	50.5(11.2)	43.9(10.2)	47.2(11.1)
FKGL	11.1(1.22)	10.7(1.15)	10.9(1.20)	10.4(1.73)	11.1(1.23)	10.8(1.53)
Syntactic Features						
NP density	.74(.20)	.72(.20)	.73(.20)	.94(.22)	.69(.13)	.81(.22)
subject density	4.72(2.79)	3.75(1.31)	4.24(2.22)	5.83(3.15)	4.73(1.62)	5.28(2.55)
Connectives						
causal connectives	46.7(29.8)	60.9(22.3)	53.8(27.1)	33.7(20.2)	36.9(15.2)	35.3(17.8)
additive connectives	9.94(9.77)	10.8(8.32)	10.4(9.01)	7.78(8.52)	19.9(10.3)	13.8(11.2)
temporal connectives	17.1(16.2)	14.2(12.4)	15.6(14.4)	6.10(10.2)	9.56(9.52)	7.83(9.92)
all connectives	133(31.7)	135(31.5)	134(31.3)	90.6(22.4)	10(22.7)	97.6(23.4)

The present analyses are in line with the preliminary analyses with regard to the insignificant effects of planning on L2 learners' performances in the cause-and-effect genre. Also, the greater structural complexity in the argumentative genre in the planning condition lend additional support to the previous findings (Ellis & Yuan, 2004; Rahimi & Zhang, 2018; Tabari, 2017) in which those in the planning condition produced more complex sentences in the argumentative genre. However, this study furthers the previous study by showing that the positive effects of planning on L2 learners' structural complexity in the argumentative genre also hold for an integrated writing task. In an integrated writing task used in the present study, L2 learners face dual demands of reading comprehension and written productions out of a source text. Thus, the L2 learners without any planning time may have their attention divided into two different tasks, which can lead to their insufficient understanding of the text. Their vague understanding of the textual elements can prompt them to excessively rely on the connectives to connect the textual elements. In fact, superficial cohesion is evident in lower-level learners' written compositions in which they overuse or misuse connectives to mask their insufficient ability to logically organize textual elements. The textual relationships are not connected smoothly without recourse to connectives; the use of which only provides "surface logicity" (Cheng & Tsang, 2021, p. 4). In contrast, those in the planning condition can pay undivided attention to each task. With enhanced understanding of the text along with adequate amount of time for writing, their writings contain fewer connectives. Also, those in the planning group appeared to be better able to support their argument by using more modifiers, evidenced in their greater NP density scores.

Discourse-level analyses

The current study includes four cohesion measures for the discourse-level analyses: argument overlap between adjacent sentences and across all sentences in referential cohesion and LSA cosine between adjacent sentences and across all sentences for semantic cohesion. The discourse-level analyses were provided in Table 5.

The statistical analyses demonstrate that the main effect of planning and the interaction effects of planning and genre are not statistically significant on argument overlap across all sentences ($F(3, 116) = 1.389, p = .241$ for the main effect of planning, $F(3, 116) = .301, p = .585$ for the interaction effect). However, main effects of planning were observable in argument overlap between adjacent sentences ($F(3, 116) = 5.620, p = .019$) and LSA cosine across all sentences ($F(3, 116) = 10.246, p = .002$). For both measures, the no planning group produced more cohesive sentences.

TABLE 5

Discourse-level Analyses on L2 Summary Writings as a Function of Planning and Genre

	Cause and effect (n = 60)			Argumentative (n = 60)		
	Planning (n = 30)	No planning (n = 30)	Total (n = 30)	Planning (n = 30)	No planning (n = 30)	Total (n = 60)
Referential Cohesion						
arg. overlap (adjacent)	.54(.27)	.62(.30)	.59(.28)	.62(.30)	.77(.21)	.69(.27)
arg. overlap (all)	.58(.24)	.74(.23)	.59(.23)	.66(.26)	.74(.23)	.70(.25)
Semantic Cohesion						
LSA cosine (adjacent)	.20(.07)	.19(.05)	.20(.06)	.25(.12)	.36(.12)	.30(.13)
LSA cosine (all)	.17(.05)	.20(.07)	.18(.06)	.27(.09)	.34(.11)	.30(.11)

The interaction effects of planning and genre were obtained on LSA cosine between adjacent sentences. The analysis on the simple main effects showed that the writings in the no planning group were more semantically cohesive only in the argumentative genre but no significant differences between the planning groups were identified in the cause-and-effect genre. The higher cohesion in the no planning condition in the argumentative genre can be explained by the learners' copying strategy. No planning group may not have enough processing time to paraphrase the source text adequately. Previous research on learners' strategies in summary writings (e.g., Keck, 2014; Kim, 2015) indicates that one of the most common and frequent strategies is copying from the source. As the source, as a model text, is higher in cohesion than students' original essays, the greater cohesion in the no planning condition is obtained from their adoption of this specific strategy. Overall, the present finding on the discourse-level features adds to past research by demonstrating that planning time may have limited effects on cohesion features. This confirms the claim that learners do not necessarily benefit from planning in all aspects of writing (Johnson, 2014; Johnson et al., 2012; Rahimi & Zhang, 2018; Rostamian et al., 2018; Tabari, 2016, 2017, 2020). The L2 learners in the current study primarily focused on the word- and sentence-level linguistic features to enhance their summary writings when additional planning time was offered.

Conclusion

The present study purports to examine the effects of pre-task planning and genre (argumentative vs. expository text) on L2 learners' integrated summary writings. To this end, this study employs a total of 22 word, sentence, and discourse-level features. According to the preliminary analyses, the source text significantly influences the learners' writing performances, especially when they are not given planning time prior to the task. Notably, the interaction of the source text and the planning variable is more prominent in the argumentative than the cause-and-effect expository genre. Consistent with the preliminary analyses, the main analyses showed that the learners were not significantly different in the

majority of the measures in the cause-and-effect expository genre, whereas there are substantial interaction effects of planning and genre in the argumentative genre. Planning overall promoted more complex language use in the word- and sentence-levels like previous studies (Rahimi & Zhang, 2018; Tabari, 2017). That is, those in the planning group displayed greater lexical complexity (as indexed by greater lexical diversity and more difficult, abstract words of lower concreteness and imageability), higher readability scores (as indicated by the lower FRE and higher FKGL scores), and greater syntactic complexity. The greater complexity, particularly in the sentence level, reflects the learners' attempts to compose more genre-specific writings. That is, they may have been motivated to use more modifiers to better support their claims. Despite the positive effects of planning in the word and sentence levels, these effects are not observed in the discourse-level. In other words, planning was shown to be less effective for enhancing their referential and semantic cohesion.

The close examination of the learners' composition in the no planning condition further reveals that the learners in no planning condition primarily drew on the copying strategy, one of the most frequent strategies in an integrated writing task (e.g., Keck, 2014; Kim, 2015). Additionally, they used more first-person pronouns, which are not regarded as a suitable pronoun for both academic genres of the study given the lowered objectivity associated with the first-person pronouns. Also, the learners without planning time included significantly more connectives in their writings than those with planning time. The overuse or misuse of connectives, frequently featured in lower-level learners' writings, can mask their poor ability to organize textual relationships in a logical way.

The findings of this study provide some pedagogical and academic implications for English education.

In terms of lexical and sentence levels, the findings show that pre-task planning was more effective in an argumentative summary than cause-and-effect summary writing. Because L2 learners do not have adequate cognitive time, they simply tend to copy words and connectives from the original text without strategizing ahead of time. Teachers should be aware that while teaching L2 reading and writing integrated tasks, students should have enough pre-task planning time so that they do not just copy or take words and phrases from the source texts. Also, this study differs from previous studies in that it investigated the influence of planning on L2 learners' performance in integrated writing, which involves reading and summarizing source texts. The current study investigating the interactions between pre-task planning and the genres of source through the analysis of two distinct source texts is differentiated from earlier research concentrating on a single genre of source text. It should be also noted that previous studies have investigated L2 students' writings in terms of CALF (complexity, accuracy, lexical sophistication, and fluency) measures relying on a paucity of indices and human raters' evaluation; however, the present study is more systematic and scientific in that some (i.e., complexity, lexical sophistication, fluency) of CALF measures were analyzed more objectively and broadly using Coh-Metrix, a multilevel language analysis tool.

Even though this study has some pedagogical and academic implications, it has inevitably some limitations. This study exclusively targeted college students majoring in English education in a Korean university, which makes it challenging to generalize the findings to other EFL learners. A future study can expand the current study by examining the writing from a broader range of EFL students. The present study also compared only the differences between the two genres (i.e., types). A follow-up study can contribute to English education by investigating more diverse types of source texts. Future research can also draw the implications of the two types of writings (i.e., integrated writing vs. independent writing) for English education by analyzing them with Coh-Metrix.

The Authors

Yoonhee Choe (first author) is an associate professor in the Department of English Education at Chongshin University. Her main interests include task-based language teaching, integrated L2 writing, English teacher education, machine translation, and technology-based L2 teaching and learning.

Department of English Education
Chongshin University
Email: yhchoi@chongshin.ac.kr

Jisu Ryu is a lecturer in the Department of English Language & Literature at Konkuk University. Her main interests include L2 reading, corpus linguistics, psycholinguistics, eyetracking, EEG, cognitive linguistics, and language analysis tools.

Department of English Language & Literature
Konkuk University
Email: jsryu0508@konkuk.ac.kr

Moongee Jeon (corresponding author) is a professor in the Department of English Language & Literature at Konkuk University. His main research interests include corpus analysis, language psychology, English education, artificial intelligence, cognitive science, and software development.

Department of English Language & Literature
Konkuk University
Email: mjeon1@konkuk.ac.kr

References

- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2). *Distributed by the Linguistic Data Consortium, University of Pennsylvania*.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839. <https://doi.org/10.1038/nrn1201>
- Cheng, A., & Tsang, A. (2020). Use and understanding of connectives: an embedded case study of ESL learners of different proficiency levels. *Language Awareness*, 1-20. <https://doi.org/10.1080/09658416.2021.1871912>
- Choe, Y. (2013). The effects of writing task types in the National English Ability Test Level 2 on Korean EFL learners' English writing performance. *Korean Journal of English Language and Linguistics*, 13(1), 161-186. <https://doi.org/10.15738/kjell.13.1.201303.161>
- Choe, Y., Ryu, J., & Jeon, M. (2020). Effects of source text's types on Korean EFL college learners' summary writing performance. *The Journal of Linguistic Science*, 95, 331-355. <https://doi.org/10.21296/jls.2020.12.95.331>
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33(4), 497-505. <https://doi.org/10.1080/14640748108400805>
- Ellis, R. (2022). Does planning before writing help? Options for pre-task planning in the teaching of writing. *ELT Journal*, 76(1), 77-87. <https://doi.org/10.1093/elt/ccab051>
- 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(1), 59-84. <https://doi.org/10.1017/s0272263104026130>

- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational researcher*, 40(5), 223-234. <https://doi.org/10.3102/0013189x11413260>
- Fazilatfar, A., Ali, M., Kasiri, F., & Nowbakht, M. (2020). The comparative effects of planning time and task conditions on the complexity, accuracy, and fluency of L2 writing by EFL learners. *Iranian Journal of Language Teaching Research*, 8(1), 93-110.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202. <https://doi.org/10.3758/bf03195564>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman. <https://doi.org/10.4324/9781315836010>
- Jeon, M., & Choe, Y. (2019). A Coh-Metrix analysis of Korean EFL learners' summary writings in the English argumentative and expository texts. *Korea Journal of English Language and Linguistics*, 19(3), 539-559. <https://doi.org/10.15738/kjell.19.3.201909.539>
- Johnson, M. D. (2014). Does planning really help?: Effectiveness of planning in L2 writing. *Journal of Second Language Teaching & Research*, 3(1), 107-118.
- 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>
- 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, 264-282. <https://doi.org/10.1016/j.jslw.2012.05.011>
- Kang, S., & Lee, J. W. (2019). Are two heads better than one? The effects of collaborative planning on L2 writing in relations to task complexity. *Journal of Second Language Writing*, 45, 61-72. <https://doi.org/10.1016/j.jslw.2019.08.001>
- Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, 25, 4-22. <https://doi.org/10.1016/j.jslw.2014.05.005>
- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing* (pp. 57-71). Lawrence Erlbaum Associates.
- Kim, N. (2015). The use of paraphrase in L2 writing with two factors: L2 proficiency and source text genre. *English Teaching*, 16(3), 71-97. <https://doi.org/10.18095/meeso.2015.16.3.04>
- Kirkpatrick, L. C., & Klein, P. D. (2009). Planning text structure as a way to improve students' writing from sources in the compare-contrast genre. *Learning and Instruction*, 19(4), 309-321. <https://doi.org/10.1016/j.learninstruc.2008.06.001>
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3), 259-284. <https://doi.org/10.1080/01638539809545028>
- Li, J. (2014). Examining genre effects on test takers' summary writing performance. *Assessing Writing*, 22, 75-90. <https://doi.org/10.1016/j.asw.2014.08.003>
- Moon, Y. (2019). Nature of planning in the writing process of L2 English essay: Focusing on brainstorming and outlining. *Foreign Languages Education*, 26(1), 109-128. <https://doi.org/10.15334/fle.2019.26.1.109>
- Oh, E., Lee, C. M., & Moon, Y. (2015). The contributions of planning, L2 linguistic knowledge and individual differences to L2 writing. *The Journal of Asia TEFL*, 12(2), 45-85.
- 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>
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21(1), 109-148. <https://doi.org/10.1017/s0272263199001047>

- Payant, C., McDonough, K., Uludag, P., & Lindberg, R. (2019). Predicting integrated writing task performance: Source comprehension, prewriting planning, and individual differences. *Journal of English for Academic Purposes*, 40, 87-97. <https://doi.org/10.1016/j.jeap.2019.06.001>
- Plakans, L. (2009). Discourse synthesis in integrated second language writing assessment. *Language Testing*, 26(4), 561-587. <https://doi.org/10.1177/0265532209340192>
- Rahimi, M., & Zhang, L. J. (2018). Effects of task complexity and planning conditions on L2 argumentative writing production. *Discourse Processes*, 55(8), 726-742. <https://doi.org/10.1080/0163853x.2017.1336042>
- Rahimpour, M., & Safarie, M. (2011). The effects of on-line and pre-task planning on descriptive writing of Iranian EFL learners. *International Journal of English Linguistics*, 1(2), 274. <https://doi.org/10.5539/ijel.v1n2p274>
- Rostamian, M., 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>
- Seyyedi, K., Ismail, M., 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>
- 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.org/10.1186/s40862-016-0015-6>
- Tabari, M. A. (2017). Investigating the effects of planning time on the complexity of L2 argumentative writing. *The Electronic Journal for English as a Second Language*, 21(1), 1-24.
- Tabari, M. A. (2020). Differential effects of strategic planning and task structure on L2 writing outcomes. *Reading & Writing Quarterly*, 36(4), 320-338. <https://doi.org/10.1080/10573569.2019.1637310>
- 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>
- Yoon, H-J. & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51(2), 275-301. <https://doi.org/10.1002/tesq.296>
- Yu, G. (2009). The shifting sands in the effects of source text summarizability on summary writing. *Assessing Writing*, 14(2), 116-137. <https://doi.org/10.1016/j.asw.2009.04.002>
- 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), 1-27. <https://doi.org/10.1093/applin/24.1.1>

(Received September 26, 2021; Revised February 20, 2022; Accepted March 18, 2022)