THE JOURNAL OF ASIA TEFL
Vol. 15, No. 4, Winter 2018, 945-958
http://dx.doi.org/10.18823/asiatefl.2018.15.4.4.945

The Journal of Asia TEFL
http://journal.asiatefl.org/

The Effects of Task Complexity and L2 Proficiency on L2 Written Performance

Jiyong Lee
University of Maryland, USA

Although extensive task-based research has focused on how the manipulation of task complexity affects L2 performance, less attention has been paid to its combined effects with those of L2 proficiency. In this study, 41 Korean learners of English carried out a written task with two levels of task complexity, operationalized as number of elements. Participants were divided into three groups, based on their level of L2 proficiency. Performance was assessed by syntactic complexity, lexical diversity, and accuracy measures. Significant interaction effects between L2 proficiency and task complexity were found on lexical diversity and accuracy measures, especially in the high-proficiency group. Findings are discussed in relation to Robinson's Cognition Hypothesis and attentional resources available for task performance.

Keywords: task complexity, L2 proficiency, L2 written performance, syntactic complexity, lexical diversity, accuracy measures, Cognition Hypothesis

Introduction

Second language research has witnessed a proliferation of studies on the effectiveness of tasks over the past two decades, and Task-Based Language Teaching (TBLT) has been adopted throughout Asian countries in institutional and classroom levels in recent years. Because the focus of language teaching has shifted from traditional teacher-centered lectures to more active student-centered learning, more attention on learner individual differences, such as second language (L2) proficiency, working memory, and motivation, is needed in order to introduce TBLT into the curriculum effectively. This paper investigates the effects of one type of individual difference, L2 proficiency, and how it interacts with those of task complexity on L2 written performance.

Task Complexity Models

Two major frameworks on task complexity and its effects on task performance have heavily influenced L2 research: Robinson's (2001, 2005, 2011) Triadic Componential Framework and Cognition Hypothesis, and Skehan's (1996, 1998, 2014) Trade-Off Hypothesis or the Limited Attentional Capacity Hypothesis. A major difference between the two models lies in their assumption of attentional resources, in that the former assumes multiple, non-competing resources, while the latter assumes a single pool of resources that learners may draw on. As a result, the two models differ in how task complexity could be operationalized, and how it has an effect on task performance.

The Triadic Componential framework specifies three components to task complexity: task complexity, task condition, and task difficulty. Referring to inherent task qualities, task complexity involves the cognitive demands of a task, which is further divided into two components: resource-directing variables that place cognitive demands on the learner, and resource-dispersing variables that place performative demands on the learner. Task condition refers to interactive demands of a task, such as participation and participant factors. Task difficulty concerns learner perceptions of task demands, which are affected by factors related to learner abilities and affective variables. Robinson claims that increases in task complexity, as opposed to task condition or task difficulty, should be the logical basis for task sequencing. The most common variables that have been manipulated in task-based research include +/- Here-and Now, +/- few elements, +/- intentional reasoning, and +/- planning time. Based on the Cognition Hypothesis, increases in task complexity along resource-directing variables will lead to greater linguistic complexity and accuracy, but less fluency. When learners carry out interactive tasks, task complexity effects will be shown in more interaction and negotiation of meaning, resulting in greater task input and modification of output.

Unlike Robinson's model, the Trade-Off Hypothesis claims that the learner's attentional capacity is limited, which leads to a trade-off effect as task complexity is increased—increased fluency may be accompanied with greater accuracy or complexity, but greater accuracy and complexity cannot occur at the same time. With the assumption of this trade-off, Skehan's model claims that balanced language development, in terms of complexity, accuracy, and fluency, can be achieved by choosing effective tasks and certain task conditions selectively. Task conditions such as familiarity of information, task interactivity, degree of structure, complex outcomes, and transformation of information can be manipulated to achieve the desired pedagogic goal.

Task Complexity and L2 Proficiency

Although the Cognition Hypothesis and the Trade-Off Hypothesis have received substantial interest by researchers and teachers, the focus of most research has been directed to task complexity effects, and to a lesser degree, task condition effects. On the other hand, the effect of L2 proficiency and its interaction with task complexity effects and/or task performance is a relatively neglected topic. In most cases, L2 proficiency is a variable that researchers attempt to control for. Only a handful of relevant studies on this topic have been found: Ishikawa (2006), Gilabert (2007), Kuiken and Vedder (2007), and Kim (2009).

Using scores on the Michigan English Placement test, Ishikawa (2006) divided 54 Japanese high school students into high- and low-proficiency groups. The two groups were further divided into two more groups depending on the task complexity condition to which they were assigned. They performed a narrative writing task, and the presence/absence of a strip cartoon during task performance, i.e., +/- Here-and-Now, determined task complexity. Performance was assessed by four measures: target-like use (TLU) of articles, S-nodes per T-unit, type-token ratio (TTR), and words per T-unit. Significant main effects of task complexity and L2 proficiency were found on TLU articles, S-nodes per T-unit, and words per T-unit, which indicated that task complexity and L2 proficiency were largely independent from each other. However, a significant task complexity × L2 proficiency interaction effect was found on TTR, in that task complexity effects were only observed in the low-proficiency group. In short, this result suggests that task complexity effects may differ depending on proficiency level.

Gilabert (2007) investigated whether learners of varying L2 proficiency levels showed different self-repair behaviors. Placement tests measuring vocabulary size was used to divide 42 Spanish learners of English into high- and low-proficiency groups. Three types of spoken tasks were employed: a narrative task, an instruction-giving task, and a decision-making task. +/- Here-and-Now, +/- few elements, and +/- intentional reasoning was used to operationalize task complexity for each task, respectively. Self-repair behavior was examined in terms of a variety of measures, and it was found that the low-proficiency group self-repaired more frequently than the high-proficiency group only in terms of error-repairs per AS-unit and ratio of error-repairs to words in the narrative task. Gilabert attributed such results to the possibility that high-proficiency learners did not encounter problems with verb morphology.

In a large-scale study with 91 Dutch learners of Italian and 76 Dutch learners of French, Kuiken and Vedder (2007) assumed that no or smaller effects of task complexity would be observed in low-proficiency learners because of the need to focus more on basic formulation processes, while high-proficiency learners were expected to show differences in performance caused by task complexity manipulations. Scores on a cloze test was used to measure English L2 proficiency, and participants carried out two written tasks, whose task complexity was manipulated along number of elements. Although researchers found significant main effects of task complexity on accuracy measures and significant main effects of L2 proficiency on accuracy and lexical variation measures for both Italian and French learners, no significant interaction effects between the two variables were found on any measure. To sum up, the study revealed that task complexity effects were not related to those of L2 proficiency.

Different findings were obtained in Kim's (2009) study of task complexity effects on language related episodes (LREs). Thirty-four ESL learners were divided into two proficiency groups based on their enrollment status in the Intensive English Program and scores on the Test of English as a Foreign Language (TOEFL). They carried out two picture narration tasks and two picture difference tasks, in which reasoning demands and number of elements were manipulated. Depending on task-type and L2 proficiency, different patterns of LRE occurrences were observed. In the case of the picture narration tasks, low-proficiency learners produced significantly more LREs when carrying out the simple task, but high-proficiency learners produced significantly more LREs in the complex task. On the other hand, in the case of the picture difference tasks, low-proficiency learners produced significantly more LREs in the complex task, while the high-proficiency learners did not show any significant difference between the simple and complex task. Kim claimed that differences in attentional resources required for task performance may have led to such results.

Considering the relative lack of consistent findings on the effects of L2 proficiency on task performance, the present study examined the combined effects of task complexity and L2 proficiency by looking into the syntactic complexity, lexical diversity, and accuracy of L2 writing, in which participants were divided into three proficiency groups.

Research Questions

The present study sought to answer the following research questions.

RQ1. Does increasing task complexity in terms of number of elements lead to changes in cognitive load?
RQ2. Does increasing task complexity in terms of number of elements lead to changes in L2 written performance?
RQ3. Do task complexity effects on L2 written performance differ according to the level of L2 proficiency?

As task complexity was manipulated in terms of number of elements, a resource-directing variable in Robinson's Cognition Hypothesis, it was hypothesized that a greater number of elements would lead to greater cognitive load, which in turn, would lead to greater syntactic complexity, lexical diversity, and accuracy. Due to mixed findings on the combined effects of task complexity and L2 proficiency, null hypotheses were formulated for the third research question.

Vol. 15, No. 4, Winter 2018, 945-958

Methodology

Participants

Forty-one learners of English as a foreign language (EFL) were recruited from a university in South Korea (18 males and 23 females). All of them were undergraduate students, the majority of whom were majoring in English Education. None of the participants had lived in an English-speaking country for more than one year, except for one participant who had lived in the United States for three years. Their ages ranged from 19 to 27 years old ($M = 22.24$, $SD = 1.95$). They were divided into three groups—low-, mid-, and high-proficiency groups—based on their scores on Brown's (1980) cloze test, which was used to measure overall English proficiency. Used for assessing vocabulary, morphosyntactic knowledge, and discourse competence, this test is considered to be a reliable and valid measure of overall English proficiency (Chrabaszcz & Jiang, 2014). There were fifty blanks in the passage, and participants were instructed to fill in the blanks with one word per blank. The exact scoring method was used. One score was given for each correct answer, yielding a final score of up to 50 points. Participants' scores were converted into z-scores, which were used to divide the participants into three groups according to their proficiency level: *low* = z-scores < -0.5, *mid* = -.05 < z-scores < 0.5, and *high* = z-score > 0.5. Table 1 provides a summary of the raw cloze test scores by proficiency levels.

TABLE 1.
Cloze Test Scores by Proficiency Groups

	N	*M*	*SD*	*Min*	*Max*
Low	12	10.58	1.78	8	13
Mid	18	15.44	1.15	14	17
High	11	21.45	3.11	18	28

Written Task

Participants carried out a written Venue Task, whose task complexity was manipulated along number of elements. They were required to choose the best venue for a special event and write a letter to the imaginary hosts. Information included a number of locations and a number of categories that had to be considered for selection. The task was closed, i.e., it had only one correct answer that participants had to find. It had a simple and complex version, which differed in the number of locations to choose from, the amount of information about the locations, and number of requirements that needed to be fulfilled in order for the participants to find the correct answer. In order to ensure that participants would not copy any English structures, all texts were provided in Korean.

Cognitive Load Measures

Self-ratings and time-on-task were used to measure cognitive load. After carrying out each task version, they answered questions on a 9-point Likert scale regarding: (1) the overall perceived difficulty of the task, (2) the level of mental effort required for task performance, and (3) the level of stress they felt during task performance.

Because the study was carried out on Enterprise Learning Management System (ELMS), an online system that logs student activity with timestamps, it was possible to measure the exact time that participants spent on the planning and writing stages of task performance. For each task version, there were two questions that participants had to answer. In the first question, all task-relevant information was provided and participants were given time to plan what to write. When they finished planning their letter, they clicked on a button that said 'I am ready to write' and then clicked 'next' to proceed to the writing stage in the second question, where they had to type their letter in a blank box. Using this method, it was possible to differentiate between time-on-planning, time-on-writing, and time-on-whole task.

Procedure

Participants met with the researcher for one session. At the beginning of the session, they took Brown's cloze test to measure their overall English proficiency. Although it was untimed, most of them took approximately 30 minutes to complete the test. After completing a language background questionnaire, they carried out the Venue Task, which had a simple and complex version. The order of the task versions was randomized. After performing each task version, participants completed a questionnaire, in which they indicated their ratings of overall perceived difficulty, mental effort required, and level of stress.

Linguistic Outcome Measures

Task performance was assessed in terms of syntactic complexity, lexical diversity, and accuracy. Syntactic complexity was measured by the number of subordinate clauses per T-unit, calculated by dividing the total number of subordinate clauses by the total number of T-units. Guiraud's Index of Richness (1954) was used to measure lexical diversity, because it is claimed to take text length into consideration. This was calculated by dividing the total number of types by the root square of the total number of tokens. Accuracy was measured in terms of the proportion of TLU of articles, calculated by dividing the total number of correctly used articles by the total number of instances that required or did not require an article.

Results

Cognitive Load Measures

Participant self-ratings and time spent on task performance were used to measure the extent to which the task versions placed a burden on cognitive processing. Descriptive statistics for these measures are provided in Table 2 and Figure 1 for self-ratings, and Table 3 and Figure 2 for time-on-task. Based on the patterns shown in the descriptive statistics, increased task complexity appeared to have a consistent effect on cognitive load, such that the complex version was perceived to be more difficult and stressful, and to require more mental effort and more time for task performance.

TABLE 2.
Means and Standard Deviations of Self-ratings

Complexity	Proficiency	Difficulty	Mental Effort	Stress	*N*
Simple	Low	2.75 (1.60)	4.08 (2.28)	3.00 (1.48)	12
	Mid	3.56 (1.79)	3.72 (1.84)	3.06 (1.73)	18
	High	2.85 (1.17)	4.36 (1.69)	2.91 (0.83)	11
Complex	Low	4.75 (1.96)	5.42 (2.19)	4.67 (2.61)	12
	Mid	4.50 (1.95)	4.83 (2.09)	3.89 (2.06)	18
	High	4.09 (1.51)	5.55 (1.37)	4.09 (1.87)	11

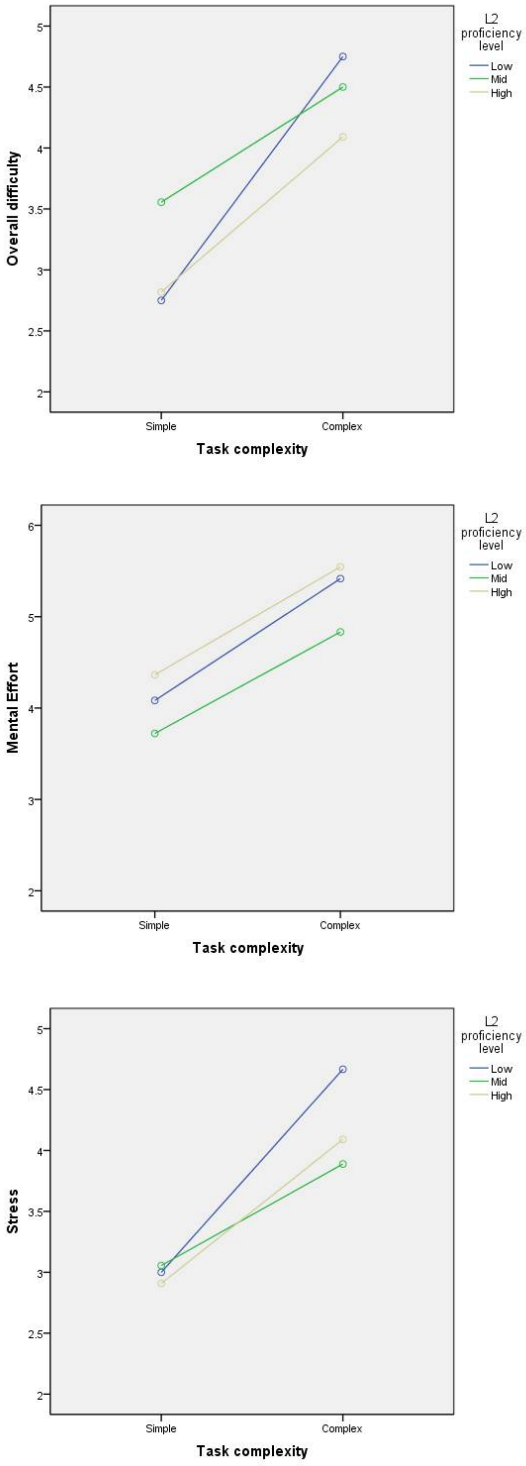

Figure 1. Self-ratings of cognitive load.

Vol. 15, No. 4, Winter 2018, 945-958

TABLE 3.
Means and Standard Deviations of Time-on-task (in seconds)

Complexity	Proficiency	Planning	Writing	Whole task	*N*
Simple	Low	66.91 (39.94)	323.00 (142.42)	398.92 (166.91)	12
	Mid	63.39 (37.24)	467.83 (228.40)	531.22 (248.68)	18
	High	60.45 (26.45)	491.09 (334.28)	551.55 (346.38)	11
Complex	Low	143.58 (173.43)	488.67 (186.40)	632.25 (246.59)	12
	Mid	128.61 (84.57)	529.06 (264.26)	657.67 (297.06)	18
	High	132.00 (74.64)	671.18 (380.26)	803.18 (445.18)	11

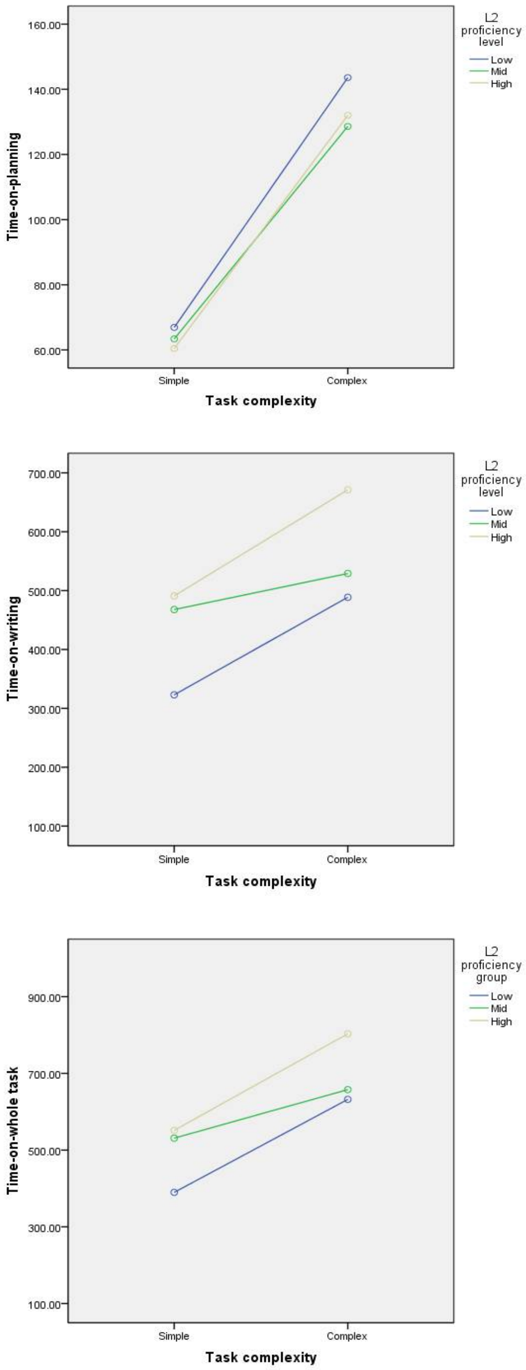

Figure 2. Time-on-task measure.

In order to see whether these trends were significant, a 2 × 3 repeated measures ANOVA was computed on the cognitive load measures, with Task complexity as the within-subjects variable with two levels (simple/complex) and L2 proficiency as the between-subjects variable with three levels (low-/mid-/high-proficiency groups). The repeated measures ANOVA was run on SPSS, with the alpha level set at .05. In line with the hypotheses of the study, it was found that increasing task complexity led to significantly higher ratings of overall perceived difficulty, mental effort required, and level of stress, $F(1, 38) = 30.916$, $p < 001$, $\eta_p{}^2 = .449$; $F(1, 38) = 21.302$, $p < .001$, $\eta_p{}^2 = .359$; and $F(1, 38) = 15.064$, $p < .001$, $\eta_p{}^2 = .284$, respectively.

Participants also spent a significantly longer time on the planning stage, writing stage, and task as a whole when performing the complex version, $F(1, 38) = 15.466$, $p < .001$, $\eta_p{}^2 = .289$; $F(1, 38) = 12.811$, $p = .001$, $\eta_p{}^2 = .252$; and $F(1, 38) = 19.371$, $p < .001$, $\eta_p{}^2 = .338$, respectively. However, no significant differences were found between the three L2 proficiency groups, $F(2, 38) = 1.617$, $p = .212$, $\eta_p{}^2 = .078$ for overall difficulty; $F(2, 38) = .067$, $p = .936$, $\eta_p{}^2 = .003$ for mental effort; $F(2, 38) = .638$, $p = .534$, $\eta_p{}^2 = .033$ for stress; $F(2, 38) = .038$, $p = .963$, $\eta_p{}^2 = .002$ for time-on-planning; $F(2, 38) = .087$, $p = .917$, $\eta_p{}^2 = .005$ for time-on-writing; and $F(2, 38) = .087$, $p = .917$, $\eta_p{}^2 = .005$ for time-on-whole task. Moreover, the interactions between task complexity and L2 proficiency were not found to be significant, $F(2, 38) = 1.617$, $p = .212$, $\eta_p{}^2 = .078$ for overall difficulty; $F(2, 38) = .067$, $p = .936$, $\eta_p{}^2 = .003$ for mental effort; $F(2, 38) = .639$, $p = .534$, $\eta_p{}^2 = .033$ for stress; $F(2, 38) = .038$, $p = .963$, $\eta_p{}^2 = .002$ for time-on-planning; $F(2, 38) = 1.123$, $p = .336$, $\eta_p{}^2 = .056$ for time-on-writing, and $F(2, 38) = .848$, $p = .436$, $\eta_p{}^2 = .043$ for time-on-whole task.

As predicted, increasing task complexity along number of elements did place a greater processing load on participants. In other words, the task version that was intended to be more complex was perceived as such. However, no significant differences were found between the three L2 proficiency groups in terms of cognitive load.

Linguistic Outcome Measures

The present study investigated L2 task performance in terms of syntactic complexity, lexical diversity, and accuracy. Table 4 and Figure 3 provide the descriptive statistics for the three linguistic outcome measures. As predicted, task complexity appeared to have a positive effect on lexical diversity, in that the complex version elicited more diverse vocabulary than the simple version. However, contrary to predictions, the complex version seemed to have negative effects on syntactic complexity and accuracy. In terms of L2 proficiency, only the high-proficiency group showed a drastic decrease in the number of subordinate clauses per T-unit when carrying out the complex task version. On the other hand, increasing task complexity appeared to have a weaker negative impact on the proportion of TLU articles for higher proficiency levels.

TABLE 4.
Means and Standard Deviations of Linguistic Outcome Measures

Complexity	Proficiency	Sub clauses per T-unit	Guiraud's Index	Proportion of TLU articles	*N*
Simple	Low	.32 (.20)	5.22 (.559)	.86 (.10)	12
	Mid	.73 (.54)	5.71 (.81)	.87 (.08)	18
	High	1.262 (1.55)	5.39 (.87)	.87 (.08)	11
Complex	Low	.40 (.21)	5.89 (.59)	.83 (.08)	12
	Mid	.73 (.37)	6.30 (.56)	.79 (.08)	18
	High	.43 (.26)	6.55 (.88)	.85 (.16)	11

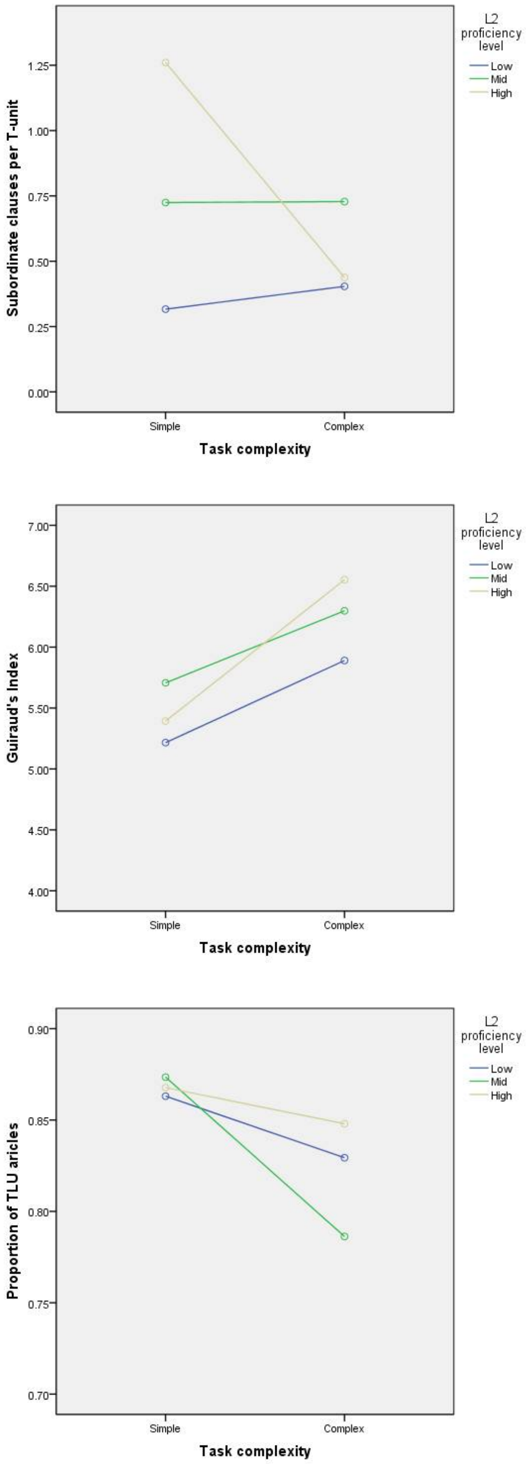

Figure 3. Linguistic outcome measures.

A 2 × 3 repeated measures ANOVA was computed on the linguistic outcome measures to identify the effects of task complexity and L2 proficiency on L2 writing, with Task complexity as the within-subjects variable with two levels (simple/complex) and L2 proficiency as the between-subjects variable with three levels (low-/mid-/high-proficiency groups). A significant interaction between Task complexity and L2 proficiency was found on the number of subordinate clauses per T-unit, $F(2, 38) = 4.315$, $p = .020$, $\eta_p^2 = .081$. The main effects of Task complexity and L2 proficiency were not found to be significant, $F(1, 38)$

= 3.350, p = .075, η_p^2 = .081; F(2, 38) = 2.993, p = .062, η_p^2 = .136, respectively. A simple effects test was computed, which revealed that 1) the high-proficiency group and the low-proficiency group showed a significant difference in number of subordinate clauses per T-unit when carrying out the simple version (p = .042), 2) the mid-proficiency group significantly outperformed the other two groups when carrying out the complex version (p = .020 when compared with the low-proficiency group, and p = .050 when compared with the high-proficiency group), and 3) the high-proficiency group used significantly fewer subordinate clauses per T-unit when carrying out the complex version (p = .005).

When a repeated measures ANOVA was computed with lexical diversity as the dependent variable, a significant main effect of Task complexity was found, F(1, 38) = 68.979, p < .001, η_p^2 = .645. A marginally significant Task complexity × L2 proficiency interaction effect was also found, F(2, 38) = 3.178, p = .053, η_p^2 = .143. On the other hand, the main effect of L2 proficiency was not found to be significant, F(2, 38) = 1.919, p = .161, η_p^2 = .092. Results of a simple effects test showed that all proficiency groups were significantly impacted by increases in task complexity (p < .001 for all comparisons), with effects being the greatest for the high-proficiency group. In other words, not only did task complexity have a positive effect on lexical diversity, it had a more beneficial effect on those with higher L2 proficiency.

Results of a repeated measures ANOVA with the proportion of TLU articles as the dependent variable revealed that Task complexity had a significant effect on accuracy, F(1, 38) = 5.705, p = .022, η_p^2 = .131. However, contrary to predictions, it was found to have a negative impact on accuracy, such that participants used fewer TLU articles when carrying out the complex task version. Furthermore, the main effect of L2 proficiency and its interaction with Task complexity were not found to be significant, F(2, 38) = .704, p = .501, η_p^2 = .036; F(2, 38) = 1.249, p = .298, η_p^2 = .062, respectively.

Discussion and Conclusion

The present study investigated whether increases in task complexity along number of elements would lead to systematic changes in cognitive load, which in turn, would lead to the desired changes in L2 written performance. It also examined whether task complexity had a differential effect on learners with varying levels of L2 proficiency.

The first research question concerned whether increasing task complexity led to the intended changes in cognitive load. Many task-based studies have overlooked this step, assuming that participants would perceive more complex tasks as such, without empirical testing. Using learner self-ratings and time-on-task, the present study obtained independent empirical evidence that increasing the number of elements in a task had a positive effect on cognitive load. In other words, regardless of L2 proficiency level, participants rated the complex version to be significantly more difficult, stressful, and to require more mental effort, and took a significantly longer time to plan their letter and write it. Such findings are consistent with those of Lee (2018), in which task complexity was operationalized as number of elements in three levels: least complex, mid-complex, and most complex. Using learner self-ratings, prospective time judgments, and a dual task method, it was found that the cognitive load of the task increased incrementally as task complexity increased. Other studies that increased task complexity along number of elements also found that the complex versions were perceived as such, obtaining evidence from learner self-ratings (see Gilabert, Baron, & Llanes, 2009; Malicka & Levkina, 2012; Sasayama, 2016).

Although learner self-ratings have been commonly used in various areas of research, time-on-task has only been introduced to task-based studies recently, and the present study is one of a very few that used it as a measure of cognitive load. Not only did time-on-task results corroborate with those regarding learner self-ratings, using this method made it possible to distinguish between time spent on planning and time spent on writing. Self-rating questionnaires usually do not include separate questions for planning and writing stages, and with the use of appropriate tools, time-on-task could be a valid and effective measure to look more closely into cognitive load.

Because number of elements is a resource-directing variable under Robinson's (2011) Triadic Componential Framework, it was hypothesized that increasing task complexity would lead to greater syntactic complexity, lexical diversity, and accuracy. Consistent with the hypotheses of the study, task complexity was found to have a significantly positive effect on lexical diversity, such that the complex version elicited more diverse vocabulary than the simple version. However, contrary to predictions, it had a negative effect on the proportion of TLU articles, such that participants produced fewer TLU articles when performing the complex version. Mixed results were found in terms of syntactic complexity, as statistical testing revealed a significant interaction between task complexity and L2 proficiency.

Even though the study did not reveal any significant main effects of L2 proficiency on L2 written performance, significant L2 proficiency × task complexity effects were found on lexical diversity and syntactic complexity measures. Not only did the mid- and low- proficiency groups benefit from task complexity effects in terms of lexical diversity, the highest gains were observed in the high-proficiency group. Although not statistically significant, the high-proficiency group also appeared to be affected negatively the *least* by task complexity in terms of the proportion of TLU articles. On the other hand, the high-proficiency group showed the greatest significant decrease in the number of subordinate clauses per T-unit when carrying out the complex task version. To summarize, while task complexity led to greater lexical diversity in all groups, especially the high-proficiency group, it also led to lower accuracy in all groups, and lower syntactic complexity in the high-proficiency group.

Several explanations can be suggested to account for the mixed findings. Although the complex task version included more elements that had to be considered, it wasn't necessary for participants to use *more* subordinate clauses when writing about the added elements. In fact, some participants used very similar paragraph and sentence structures for both the simple and complex task versions, which may have led to the non-significant effect of task complexity on syntactic complexity. As evidenced by the low average self-ratings on the simple task version (3.12 for difficulty, 4 for mental effort, and 3 for stress) and the short average planning time (63.63 seconds), the simple version appeared to be quite easy for the participants. Because it was so easy to find the correct answer, the high-proficiency group may have had more attentional resources available to allocate to syntactic structure while performing the simple version. However, when carrying out the complex version, they may have focused more on content, vocabulary and/or accuracy, resulting in tradeoffs between syntactic complexity and lexical diversity, and between syntactic complexity and accuracy. In terms of accuracy, article use is one of the most difficult features of English to acquire by Korean participants because Korean lacks an article system, and the participants' limited proficiency in English may have made it difficult for lower proficiency groups to attend to articles when their focus was on incorporating the added elements to their writing when performing the complex task version.

Table 5 shows an example of a participant from a high-proficiency group, whose number of subordinate clauses per T-unit decreased from 1.25 to .81, while GI and proportion of TLU articles increased from 5.83 to 7.48 and .81 to .85, respectively, as task complexity increased. Although they were very detailed in their explanations on why they thought a certain venue was the most appropriate place to hold a birthday party and why others were not, they used fewer subordinate clauses in the complex version and instead used more coordinating conjunctions such as *and* or *but*. This led to a decreased value in syntactic complexity. On the other hand, because they elaborated on all of the details that needed to be explained, they used more words in their letter in the complex version, resulting in a higher GI.

Vol. 15, No. 4, Winter 2018, 945-958

TABLE 5.
Example of Venue Task letter

L2	Simple	Complex
High-proficiency	I would recommend Brian and Alice [to hold their mother's 70th birthday party in The Castle], [because The Castle satisfies all the requirements]. First of all, the meal budget per a person is $45, [which does not exceed $50.] Secondly, the room's capability of accommodation is 40 and it adequately meets their expectation. Lastly, [considering [Brian and Alice want their party place [to be quiet]]], quiet music of The Castle will fit their needs. In case of The Square, it would not meet Brian and Alice's needs [because the meal budget per person exceeds $50 by $15, and the capability of room accommodation is lower than their expectation]. The Royal is even worse [because it has got all problems [that The Square has]]. Moreover, the live music of the place will not ensure quiet atmosphere of the party place.	[Taking Liam and Kate's considerations into account,] The Tower would be the best place for the 80th birthday party [to be held]. The Tower meets their meal budget expectation, provides free parking space, has a room capable [of accommodating] 30 people, limits period of time for the room up to 3 hours, and they have visual projection equipment readily available. The live piano [being played in the party room] could be an obstacle, but piano music is usually less noisy compared to band musics. Other party places have at least one problem each. The Lunchroom meets all needs but parking place. The Lighthouse does not have its own parking place and visual equipment available, they cannot hold more than 20 people in the room, and the time for room usage is not limited at all. The Springs does not limit the time, and does not have visual equipment either. And also, they play band musics in the room [which might not be considered as a quiet atmosphere.]

Note. Square brackets indicate subordinate clauses.
Underlined words indicate coordinating conjunctions in main clauses.

Combining the findings regarding cognitive load measures and linguistic outcome measures, it could be claimed that increasing task complexity along number of elements does lead to greater cognitive load, which does not necessarily lead to better task performance. Thus, the study obtained evidence in support of Robinson's Cognition Hypothesis only in terms of lexical diversity. The significant interactions between task complexity and L2 proficiency indicate that the effects of task complexity seem to be greater in high-proficiency learners. However, the findings regarding the negative effect of task complexity on high-proficiency learners' use of subordinate clause per T-unit are tentative, given the small number of participants in the high-proficiency group ($N = 11$).

Limitations of the study should be acknowledged, which include small number of participants, use of one type of written task, and only two levels of task complexity. For future investigations, more elements could be manipulated in order to increase task complexity to a substantial degree. Considering that the self-ratings were based on a 9-point Likert scale, the ratings of overall difficulty, mental effort, and level of stress in the present study were not very high on the complex task version: 4.46, 5.20, and 4.17, respectively, averaged across proficiency levels. Different results may have been obtained if the complex version were more cognitively challenging. Future studies could also employ more diverse task-types to generalize findings.

Acknowledgement

I would like to express my gratitude to Dr. Michael Long for his guidance valuable suggestions. I would also like to thank Drs. Robert DeKeyser and Steven Ross for their support. My gratitude also extends to the anonymous reviewers.

Vol. 15, No. 4, Winter 2018, 945-958

The Author

Jiyong Lee completed her Ph.D. in Second Language Acquisition at the University of Maryland, USA in December 2018.

Second Language Acquisition Program
University of Maryland
1102 Francis Scott Key Hall
College Park, MD 200742
Email: jlee0123@umd.edu

References

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, *64*(3), 311-317.

Chrabaszcz, A., & Jiang, N. (2014). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, *30*(3), 351-379.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and (+/- here-and-now): Effects on L2 oral production. In M. d. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44-68). Tonawanda, NY: Multilingual Matters Ltd.

Gilabert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching, 47*(3-4), 367-395.

Guiraud, P. (1954). *Les Charactères Statistiques du Vocabulaire. Essai de méthodologie*. Paris: Presses Universitaires de France.

Ishikawa, T. (2006). The effect of task complexity and language proficiency on task-based language performance. *The Journal of AsiaTEFL*, *3*(4), 193-225.

Kim, Y. (2009). The effects of task complexity on learner–learner interaction. *System*, *37*(2), 254-268.

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching, 45*(3), 261-284.

Lee, J. (2018). Task complexity, cognitive load, and l1 speech. *Applied Linguistics, 39*(3). doi:10.1093/applin/amy011

Malicka, A., & Levkina, M. (2012). Measuring task complexity: Does L2 proficiency matter. In A. Shehadeh & C. A. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 43-66). Philadelphia, PA: John Benjamins Publishing.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). New York, NY: Cambridge University Press.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*(1), 1-32.

Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 3-37). Philadelphia, PA: John Benjamins Publishing.

Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of cognitive task complexity. *The Modern Language Journal, 100*(1), 231-254.

Vol. 15, No. 4, Winter 2018, 945-958

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38-62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2014). *Processing perspectives on task performance*. Philadelphia, PA: John Benjamins Publishing.