Effects of Task Complexity on English Argumentative Writing

Hyejin Cho
(Sookmyung Women’s University)

*English Teaching, 70*(2), 107-131.

The purpose of this study is to examine the effects of task complexity on the quality of L2 learners’ argumentative writing using both global measures of complexity, accuracy, and fluency (CAF) and task-specific measures, namely, for conjunctions. A group of 110 Korean high school students in South Korea performed either a simple or complex argumentative writing task. Task complexity was manipulated by +/-reasoning demands and +/-few elements. A set of 110 argumentative essays were analyzed on 6 global measures of CAF and 2 task-specific measures. The results showed that task complexity affected the fluency of the argumentative writings, in that the complex task group produced more fluent writings than the simple group. However, task complexity did not affect accuracy or syntactic complexity of the argumentative writings. In the task-specific measures, task complexity affected neither frequency nor target-like use of conjunctions. These results have pedagogical implications for task design to help learners develop their L2 proficiency.

**Key words:** task complexity, global measure, task-specific measure, argumentative writing

1. INTRODUCTION

In the 1980s, the foci of second language acquisition (SLA) research were twofold: to describe natural characteristics of language acquisition and to test theory-based research hypotheses (Ellis, 2001). The emergence of theory-based research amplified task-oriented empirical studies, and Task-Based Language Teaching (TBLT) began to attract attention from language educators (Kuiken & Vedder, 2008). One of the research topics in TBLT is the effects of task complexity on learners’ L2 performance, with task complexity being determined by the cognitive demands imposed by task factors on learners (Robinson,
This interest in task complexity is because it is widely known that task designs may affect language learning processes (Shin, 2012). There are two major hypotheses concerning with task complexity: the Cognitive Hypothesis (Robinson, 2001a, 2001b), and the Trade-Off Hypothesis based on the Limited Attentional Capacity Model proposed by Skehan and Foster (1999, 2001). Robinson (1995) assumes that cognitively more demanding tasks help learners produce more accurate and more complex language production.

The Cognitive Hypothesis postulates not every complex task necessarily induces trade-off effects among different dimensions of L2 performance because different processes may draw on various attentional pools (Robinson, 2007). On the other hand, Skehan and Foster (1997) argue that learners experience trade-off effects among complexity, accuracy, fluency (CAF) of L2 production when cognitively demanding tasks are given. That is, when learners pay more attention to meaning, relatively less attention would be paid to linguistic forms during language production.

Based on these hypotheses, a number of empirical studies have investigated the effects of task complexity on L2 oral production in terms of complexity, accuracy, and fluency (Gilabert, 2005; Lee, 2002; Michel, 2013; Revesz, 2011; Robinson, 2005, 2007). However, relatively few studies have focused on the influence of task complexity on the quality of written discourse (Farahani, 2011; Hosseini & Rahimpour, 2010; Kormos & Trebits, 2012; Kuiken & Vedder, 2008). In addition, no conclusive effect of task complexity has been found for L2 oral and writing performance. As one explanation for the inconclusive empirical results, researchers have pointed out several limitations of using global measures of CAF on second language production (Michel, 2013; Robinson, 2005, 2007). In particular, Robinson, Cadierno, and Shirai (2009) claim that the traditional global measures of CAF on language production are general to examine the effects of task complexity on the quality of L2 performance, and supplementary task-specific measures are required to detect specific task effects.

Motivated by the limitations of the previous research, the study is designed to discover how cognitive complexity differently affects the qualities of learners’ argumentative L2 writing by employing both global measures of CAF and conjunctions. In the present study, task complexity is manipulated by the amount of reasoning demands and the number of elements. The reason why conjunctions function as a task-specific measure is that argumentative tasks manipulated along reasoning demands push learners to support their opinions by using conjunctions. Specifically, the following research questions guided this study:

1. What are the effects of task complexity on CAF of L2 written performance measured by global measures of CAF?
2. What are the effects of task complexity on the use of conjunctions measured by task-specific measures?

2. LITERATURE REVIEW

2.1. Cognition Hypothesis

Robinson (2001a, 2001b) postulates that human attentional resources are unlimited when learners are processing information, so he assumes that learners are able to attend to multiple linguistic aspects and process multiple aspects without trade-off effects in their language production. In order to examine the relationship between cognitive demands and task factors, Robinson (2001a) firstly defines task complexity as following:

The cognitive task features which can be manipulated either to increase or decrease cognitive demands placed on the learners when they perform a task and the result of the attentional, memory, reasoning and other information processing demands imposed by the structure of the task on the language learner. (p. 29)

Moreover, Robinson (2005) developed the Triadic Componential Framework. In this framework, task complexity consists of two dimensions: resource-directing and resource-dispersing dimensions. Resource-directing factors refer to variables that place cognitive and conceptual demands on learners while learners are performing a task such as +/−here-and-now, +/−few elements, and +/−no reasoning demands; the +feature is associated with a simple task and the −feature is associated with a complex task. Meanwhile, resource-dispersing factors refer to variables that place performative and procedural demands while a task is performed such as +/−planning time, +/−single task, and +/−prior knowledge (Robinson, 2005).

Based on these concerns, Robinson (2001a) claims that increasing task complexity along the resource-directing dimension results in more accurate and complex language production, because learners would pay more attention to the conceptual demands of a given task when they are given cognitively demanding tasks, and thus they process language more deliberately. However, the increase of task complexity along resource-directing dimensions would lead to less fluent production, because resource-directing factors demand more cognitive attention from learners. As a result, these resource-directing factors disturb fluency. Robinson (2001a, 2007) also argues that though complexity and accuracy of L2 speaking performance are related, fluency stands in contrast to accuracy and complexity. On the other hand, Robinson (2001a, 2001b) claims that along the
resource-dispersing dimension in terms of performative and procedural demands, increasing task complexity would lead to less fluent, less accurate, and less complex L2 production, because complex tasks overburden learners. This present study focuses on the resource-directing dimension.

2.2. The Limited Attentional Capacity Model

Even though both Robinson (2001a, 2001b) and Skehan and Foster (2001) are concerned with task complexity in second language production in terms of attentional allocation, their expectations and theoretical backgrounds differ. Skehan and Foster postulate that humans have limited abilities to process information at once, with their attention being split among different aspects of L2 performance. Thus, trade-off effects among different aspects of information should be found when humans are processing cognitively demanding information. Focusing on one aspect causes other aspects to suffer from a lack of attentional resources. As a result, only those performance aspects receiving sufficient attention from learners will be successfully processed when learners are performing a cognitively demanding task, while other aspects which do not receive sufficient attention will not be processed well. Skehan and Foster specifically claim that the initial distinction of performance dimensions is a distinction between a focus on form and a focus on fluency (meaning). Then the form is further broken down into accuracy and complexity. In addition, Skehan and Foster mention that “Prioritisation or predisposition (or both) seem to orient performance towards one (or two) of the three areas theorized to be important” (p. 193).

Similar to the Limited Attentional Capacity Model, VanPatten (1990) postulates that language learners have limited attentional capacities based on the information processing models and points out that learners tend to pay their attention to a certain aspect of language production, which is prioritized by meaning and communicative needs of a given task. That is, when learners are performing cognitively demanding tasks, they may experience a lack of attentional capacities caused by humans’ limited attentional resources. Then they inevitably reach a crossroad where they have to decide which dimension(s) they focus on more in order to complete a given task. As a result, learners experience trade-off effects among complexity, accuracy, and fluency in L2 performance. In other words, they may not be able to focus on all of these dimensions.

All things considered, Skehan and Foster (2001) assume that when complex tasks burden cognitive demands on learners during task performance, learners may experience trade-off effects among complexity, accuracy, and fluency of L2 production because humans have a limited cognitive capacity, which leads to trade-off effects among three dimensions of language production.
2.3. Task Complexity in Written Production

In contrast to a number of the early empirical studies simply focusing on effects of task complexity on L2 speaking performance, recent studies have started to examine the effects on L2 writing as well in order to investigate the theoretical validity of those two hypotheses in a written mode. Kormos and Trebits (2012) point out the need to distinguish L2 speaking performance from written production in task complexity. In their study, equivalent task materials were given to participants who were asked to perform an oral task and a writing task to measure the effect of production modes on task complexity. The results showed that the participants used more varied vocabulary in their writings than in speaking performance. Also, a higher ratio of error-free clauses was found in the written mode than the oral mode. It would appear that each performance mode had different characteristics, possibly affecting the effects of task complexity on L2 production. The most important pedagogical implication in this study is that teachers may need to choose an appropriate production mode depending on learner variables and the purpose of their language courses.

Motivated by the difference in language production modes, several studies were conducted. Ishikawa (2006) manipulated task complexity along +/–here-and-now, a resource-directing dimension and measured fluency of writing performance by words per T-unit. The results showed that the complex group outperformed the simple group in fluency of their performance. Moreover, regardless of learners’ proficiency levels, the positive effects of increasing task complexity along resource-directing dimensions on fluency of the writings were found. Hosseini and Rahimpour (2010) showed that participants who completed a there-and-then task as a complex task produced more fluent writings than a simple task group. They explain that the results may be attributable to the assumption that learners have limited attentional capacity, in accordance with Skehan and Foster’s hypothesis, so it would appear that there was a trade-off effect between fluency and form when learners performed a complex task.

Salimi, Dadaspour, and Asadollahfam (2011) also show that a complex task led to more fluent written performance than a simple task. They argue that this result could be attributed to the influence of task complexity along resource-directing dimensions on language learning or L2 development by attracting learners’ attention to the fluency dimension. In sum, it is found that complex tasks facilitated learners’ production of more fluent writings than simple tasks. On the other hand, the effects of task complexity on complexity and/or accuracy were not supported in these studies. However, few empirical studies on task complexity in written modes have been conducted, so it is difficult to reach a solid conclusion yet. Thus, the present study examines effects of task complexity on L2 writings, narrowed down to argumentative writing tasks.
2.4. Using Task-Specific Measure

2.4.1. Empirical studies using task-specific measures

Similar to the growing interest in the effects of task complexity on L2 writing performance, recent studies have given attention to more subtle effects of task complexity with task-specific measures as a way of overcoming the limitations of global measures. The widespread usage of global measures of CAF may be problematic, given that many studies used one measure of each dimension of CAF, respectively, and that it is difficult to examine all of the sub-constructs of each dimension of CAF (Michel, 2013). In this context, Robinson (2005, 2007) points out the insufficiency of these general measures and the necessity of using task-specific measures to more sensitively examine the effects of task complexity on accuracy and complexity of learners’ L2 oral production. According to Robinson (2005), task-specific measures are “inter language-sensitive and cognitive linguistic measures of production” (pp. 21-22).

Based on this assumption, Robinson (2007) employed both global measures of CAF and task-specific measures, psychological state terms and complex syntactic structures to analyze the same data. As a result, the complex task led to more accurate and complex oral production when the data were analyzed with task-specific measures, whereas the global measures did not detect the effects of task complexity. William and Graumann (2009) showed that task-specific measures operationalized as conjoined clauses and elaborated noun phrases in this study clearly captured impacts of task complexity on oral production than global measures. Revesz (2011) examined the effects of task complexity on oral production. The results showed that task-specific measures detected statistically significant effects of task complexity on L2 oral production while global measures of CAF did not. However, compared to the amount of interest in task-specific measures of L2 learners’ speaking performance, little attention has been given to L2 writing products. Thus, it may be meaningful to examine effects of task complexity on L2 writing by employing task-specific measures. Following Robin’s (2005, 2007) framework the present study employs conjunctions to investigate the effects of task complexity, focusing on the complexity and accuracy of the use of conjunctions.

2.4.2. Conjunctions as task-specific measures for argumentative writing tasks

Since each task design consists of different dimensions of task complexity, task-specific measures should be defined differently based on the characteristics of a particular task. As the present study employs English argumentative writing tasks, it is necessary to use certain task-specific measures relevant to argumentative writing. There have been
empirical studies on the effects of task complexity which operationalize conjunctions as task-specific measures in argumentative writing tasks. Robinson (2005) specifically claims that +/–reasoning demands accompany subordinating conjunctions, the use of psychological, cognitive state verbs, and complex complementation, because tasks manipulated according to +/–reasoning demand articulate reasons. Robinson and Gilabert (2007) also conducted a study with tasks manipulated according to +/–reasoning demands, and operationalized psychological state terms and more conjoined phrases as task-specific measures for his study based on the same rationale. William and Graumann (2009) operationalized conjoined clauses and elaborated noun phrases as task-specific measures and examined the effects of task complexity according to +/–reasoning demands on L2 oral production. The rationale for these conjoined clauses as task-specific measures is that tasks manipulated according to +/–reasoning demands coordinate more conjunctions and adverbial clauses than relatively simple tasks, which require less cognition from learners. In addition, Revesz (2011) claims that conjoined clauses are relevant to the +/–reasoning dimension as task-specific measures, because tasks manipulated according to +/–reasoning dimension force learners to justify their opinions and present arguments in order to persuade others. Michel (2013) assumes that argumentative tasks may lead to greater and more varied usage of conjunctions based on a considerable number of empirical studies which have examined the association between argumentative writing and the use of conjunctions (Byrnes, Maxim, & Norris, 2010; Christie & Derewianka, 2008; Schleppegrell, 2004). Based on these theoretical concerns, the present study employs conjunctions as a task-specific measure in order to examine whether task complexity has a significant influence on the use of conjunctions.

3. METHODOLOGY

3.1. Participants

Participants in the present study were 110 Korean learners of English at a Christian alternative high school in a city located in the middle of South Korea. They have learned English for more than six years since they were in the 3rd grade of elementary school by taking the compulsory English education program in South Korea. Most of the students have never gone abroad, with a few having lived in English-speaking counties for less than two years. Of these 110 students, 44 were females and 66 were male students. In terms of the grades, 56 students were in the 11th grade and 54 students were in the 12th grade. The English proficiency of the participants ranged from beginners to advanced levels; 9% were advanced learners qualifying as Excellent to very good, 57% were intermediate learners
qualifying as *Good to average*, and 34% were beginners qualifying as *Fair to poor*, according to Hedgcock and Lefkowitz’s (1992) scale.

3.2. Tasks

To choose a topic for the tasks, the researcher asked several of the participants’ teachers certain questions about their students’ common interests. The answers they gave led the researcher into designing experimental tasks about dormitory life. Since all the participants live in a dormitory on campus, choosing the best roommates would likely be an interesting and familiar topic for them.

Motivated by Michel’s (2013) tasks, the common format of simple and complex version is an argumentative task which asks learners to choose the best roommates based on the given profiles for each candidate and to write at least three reasons for their own decision in English. More specifically, the simple task asks learners to choose one couple out of 4 candidates, each marked by 4 properties: hobbies, personality, studying style, and sleeping patterns (see Appendix A). The complex task asks learners to choose 2 couples out of 6 candidates, each marked by 6 properties: hobbies, personality, studying style, sleeping pattern, favorite subjects, and cleanliness (see Appendix B). These properties were chosen based on common aspects about which the students were concerned when choosing their roommates from some of the students’ comments on dorm life.

Task complexity was manipulated by two resource-directing factors: +/-no reasoning demands and +/-few elements. The combination of these factors maximizes the difference between complex and simple tasks according to cognitive demands. The +/-no reasoning demands are marked by the number of properties for each candidate, and +/-few elements are represented by the number of candidates. The complex task is characterized by –no reasoning demands and –few elements, since it has 2 more properties for each candidate and 2 more candidates than the simple task. On the other hand, the simple task is marked by +no reasoning demands and +few elements. The rationale behind the manipulation of task complexity is that tasks with more reasoning demands and a greater number of elements are considered to be cognitively more complex tasks than those with less reasoning demands and few elements (Kuiken & Vedder, 2007; Revesz, 2011).

3.3. Procedures

Originally there were 122 students who initially participated in this study. To evaluate the students’ English writing competence, they were asked to write a short essay about one of the most exciting moments in their lives for 30 minutes during their regular English classes in the middle of March, 2014. All the short essays were rated by two native
speakers of English who were English teachers at a high school. They graded the short English essays based on the rubric designed by Hedgcock and Lefkowitz (1992), as it is a widely quoted and used rubric for assessing L2 writings. Interrater reliability was significant at a .05 significance level for grading the English essays ($r = .76$). Finally, the present study excluded 12 students whose writings contained less than 3 complete English sentences, because to examine CAF of written performance, incomplete sentence structures would be problematic. Thus, only 110 students were assigned to either the complex or the simple task.

Based on the average score from the two raters for each short essay, 110 students were divided into two groups, which displayed the same component ratio for English writing proficiency levels. As shown in Table 1, the homogeneity of these two groups in English writing proficiency levels was proved through an independent samples $t$-test using the mean score of each participant’s essay. At a .05 significance level, these two experimental groups were shown to be homogeneous ($p = .985$).

### TABLE 1

<table>
<thead>
<tr>
<th>Group</th>
<th>$N$</th>
<th>Mean</th>
<th>$SD$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>55</td>
<td>73.52</td>
<td>10.34</td>
<td>.018</td>
<td>.985</td>
</tr>
<tr>
<td>Complex</td>
<td>55</td>
<td>73.48</td>
<td>10.30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Two weeks later, two homogeneous groups were assigned to either the simple task or the complex task. Each task was allotted 40 minutes for completion. The participants were not allowed to use a dictionary or receive any help from their friends or teachers. All the participants fully understood what they had to do after the researcher gave guidelines. Except for the complexity of the tasks, all 110 students were under the same task performance conditions. Then the 110 writing samples were analyzed according to global measures of CAF and task-specific measures, respectively.

### 3.4. Measures

3.4.1. The definitions of production units

For measuring and analyzing complexity, accuracy, and fluency of written performance, Lu (2010) synthesized the definitions of syntactic units and redefined the terms to facilitate the analysis of written performance. The present study adapted each definition of terms redefined by Lu: sentences, clauses, dependent clauses, and T-unit. To be more specific, a T-unit refers to an independent clause and its dependent clauses.
3.4.2. Global measures of CAF

All the 110 writings were measured in three dimensions: complexity, accuracy, and fluency. The learners’ proficiency levels should be considered when choosing measures for accuracy. It was difficult to find any error-free T-units from the beginner level learners’ performance, so the number of error-free T-units and error-free T-units per T-unit are more suitable for intermediate and advanced levels, whereas the number of errors per T-unit can suitably measure the accuracy of writings written by beginning level learners (Kuiken, Mos, & Vedder, 2005; Kuiken & Vedder, 2007). Since the participants’ proficiency levels in the present study ranged from beginner levels to advanced levels, both error-free T-units per T-unit and the number of errors per T-unit were used as measures of accuracy. Moreover, these two measures are two out of three widely used measures for accuracy: the number of error-free T-units, error-free T-units per T-unit, and the number of errors per T-unit (Wolfe-Quintero, Inagaki, & Kim, 1998).

The measures of syntactic complexity are both clauses per T-unit, which is one of the subordinate clause ratios, and T-units per sentence, which is one of sentence coordination ratios. This is because novice level learners tend to rely on coordination clauses, while intermediate or upper-intermediate learners more frequently use subordinate clauses (Hwang, 2012; Wolfe-Quintero et al., 1998). For fluency, frequencies were not suitable measures of writing fluency, given the relative brevity of the tasks given to the participants. In short, fluency ratio measures, not frequency measures, were suitable for this study. Thus, fluency was measured by both the number of words per T-unit and the number of words per sentence. These two measures of fluency are widely used fluency ratio measures for L2 written performance (Hwang, 2012; Wolfe-Quintero et al., 1998).

3.4.3. Task-specific measures

In the present study, two task-specific measures are used: frequency of conjunctions as a complexity measure and target-like use of conjunctions as an accuracy measure. Frequency is defined as the number of conjunctions per 100 words. This definition is taken from Michel’s (2013) study, exploring the use of Dutch conjunctions in simple and complex argumentative tasks along with reasoning demands. Target-like use of conjunctions is the number of accurately supplied conjunctions in obligatory contexts per number of obligatory contexts plus non-obligatory contexts with inappropriate usage. This definition is taken from Gilabert’s (2005) formula, originally defining accuracy for the use of articles. Gilabert defined target-like use of articles as the number of accurately supplied articles in obligatory contexts per number of obligatory contexts plus non-obligatory contexts with inappropriate suppliance. The frequency of particular linguistic features is regarded as one
measure of complexity in a number of empirical studies (Ellis & Yuan, 2005; Michel, 2013; Revesz, 2011; Robinson, 2007; Robinson et al., 2009). Target-like use of conjunctions examines how accurately and appropriately conjunctions are used in the learners’ writing performance.

3.5. Analyses

3.5.1. Measuring global CAF

Two native speakers of English who have taught English for more than 2 years in a Korean high school marked all morphological, syntactic, and semantic errors in the writing samples. To maintain marking objectivity, the present study adapted Polio’s (1997) guidelines for analyzing linguistic errors, a comprehensive description to achieve inter-rater reliabilities by identifying linguistic errors. Most of the errors on which these two raters marked were the same. However, when there was a disagreement in marking, the raters and the researcher discussed whether to regard the apparent error as an error. For each disagreement, all the raters and the researcher reached an agreement based on the guidelines.

As a result, the participants made many errors in their usage of English articles, spelling, punctuation, and capitalization. This is consistent with the fact that 91 percent of the participants in the present study were either (1) beginners who tended to make frequent grammar errors in agreement, number, tense, or articles, and frequent mechanical errors such as spelling, punctuation, and capitalization; or (2) intermediate learners of English who made the aforementioned errors occasionally, according to Hedgcock and Lefkowitz’s (1992) rubric. Also, L2 learners tend to acquire articles in the later stages of their L2 morpheme acquisition (Dulay & Burt, 1973; Hakuta, 1976).

If all the article and mechanical errors the participants made had been counted as errors in the present study, it would have affected the validity of error-free T-units; few error-free T-units would have remained. Many other empirical studies on objective measures of beginner or intermediate level writings have not counted spelling, punctuation, and capitalization mistakes based on the same rationale of the present study (Bardovi-Harlig & Bofman, 1989; Casanave, 1994; Ishikawa, 1995). Thus, the present study considers errors as incorrect use of grammar other than articles and mechanical errors, which are not directly related to meaning.

For analysis of syntactic complexity and fluency, the number of production units such as T-units, clauses, words, and sentences was calculated by two graduate students who had majored in English Linguistics. They also followed the guidelines designed by Polio (1997), and each rater counted the number of words, T-units, clauses, and sentences in each
argumentative writing sample. Then they held a discussion when they did not reach an agreement on the number of the production units.

To measure syntactic complexity, three production units were used in the present study: T-units, clauses, and sentences. The number of clauses was divided by the total number of T-units, and the number of T-units was divided by the total number of sentences based on the number of each production unit. To measure fluency, the number of words was divided by the number of T-units and by the number of total sentences. The number of words or sentences itself is not suitable measures of writing fluency, given the relative brevity of the tasks given to the participants. In short, fluency ratio measures, not frequencies, were deemed suitable for this study. Thus three production units as elements of two widely used fluency ratio measures were used: T-units, sentences, and words.

3.5.2. Measuring the use of conjunctions

Conjunctions as task-specific measures in the present study should be clearly defined, since the definition affects the validity of the study. Halliday and Hasan (1976) established their taxonomy of conjunctions which L2 learners use in their writings to examine the L2 developmental stages, and categorized four dimensions of conjunctive relations: additive, adversative, causal, and temporal conjunctive relations. In their study, the definition of conjunctions includes both conjunctions which are followed by finite clauses, and connectors, which do not have to be followed by finite clauses. However, in this present study, conjunctions are limited to conjunctions followed by finite clauses based on the taxonomy. Table 2 shows the conjunctions examined in the present study.

<table>
<thead>
<tr>
<th>Conjunctive Relations</th>
<th>Conjunctions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive</td>
<td>and, or, and also, nor</td>
</tr>
<tr>
<td>Adversative</td>
<td>though, but, although, even though, even if</td>
</tr>
<tr>
<td>Causal</td>
<td>so, so that, if, as, because, since</td>
</tr>
<tr>
<td>Temporal</td>
<td>when, until, as, after, since, before, while, once</td>
</tr>
</tbody>
</table>

To measure frequency, the writing samples of each group were first screened for all the selected conjunctions using AntConc, a freeware concordancer. As a result, the simple group used 7 types of conjunctions: *and, so, because, but, if, when, and so that*, but one participant in the simple group used *so that* twice, so mainly only 6 conjunctions were used by most of the participants. On the other hand, the learners in the complex group used 12 different types of conjunctions: *and, so, because, but, if, when, as, or, although, though, until, and so that*. 
However, AntConc did not distinguish conjunctions followed by a clause from conjunctions connecting words and phrases, so the researcher distinguished conjunctions as connectors between words and phrases from conjunctions which connect finite clauses. The researcher inspected all of the use of conjunctions which are potentially regarded as both conjunctions followed by finite clauses and connectors between words and phrases. Also, the researcher examined some conjunctions which function not only as conjunctions but also other lexical classes.

Finally, in the simple group’s performance, the number of 7 types of conjunctions followed by finite clauses was counted. Likewise, the number of the 12 conjunctions used in the complex group’s performance was also counted, respectively. Then to prevent a bias for text length, the raw number of conjunctions used by each group was converted into frequencies per 100 words, respectively.

To measure target-like use of conjunctions as an accuracy measure in the present study, the two English-speaking raters also marked the inappropriate use of conjunctions and obligatory contexts where learners failed to use a needed conjunction in the 110 writings. Both raters discussed whether an error noted by a single rater should be counted as one. In this case, the researcher joined in the discussion to reach agreement, given that conjunction errors are not as straightforward as morphological, syntactic, and semantic errors, and that learners could use a variety of possible conjunctions to express the same meaning. An agreement was finally reached for each problematic case. The resulting data was compiled and used to calculate each group’s target-like use of conjunctions, respectively.

3.5.3. Statistical analysis

The Statistical Package for Social Science (SPSS) was used to analyze the data in the present study. All the statistical analyses were conducted at a .05 significance level. In order to investigate statistically significant differences between the complex task group’s writing performance and the simple task group’s performance, t-tests were conducted for each of the six measures of CAF and two task-specific measures on writing samples.

4. RESULTS AND DISCUSSION

4.1. Global Measures

4.1.1. Results of complexity, accuracy, and fluency

In terms of complexity, the complex group produced syntactically more complex
linguistic structures than the simple group. The complex group produced more clauses per T-unit \( (M = 1.34) \) than the simple group \( (M = 1.27) \). Also, the complex group produced more T-units per sentence \( (M = 1.20) \) than the simple group \( (M = 1.15) \), as shown in Table 3. However, no statistically significant difference between the complex group and the simple group was found in both measures of syntactic complexity \( (C/T; t = -1.42, p > .05; \) and \( T/S; t = -0.88, p > .05, \) respectively). Increasing task complexity along resource-directing dimensions did not affect syntactic complexity of the participants’ argumentative writings.

**TABLE 3**

<table>
<thead>
<tr>
<th>Statistics on Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>C/T</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>T/S</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Note. T = total number of T-units; C = total number of clauses; S = total number of sentences*

**TABLE 4**

<table>
<thead>
<tr>
<th>Statistics on Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>EFT/T</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>E/T</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Note. EFT = total number of error-free T-units; T = total number of T-units; E = total number of errors*

**TABLE 5**

<table>
<thead>
<tr>
<th>Statistics on Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>W/T</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>W/S</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Note. W = total number of words; T = total number of T-units; S = total number of sentences
* \( p < 0.05 \)

For accuracy, as shown in Table 4, the complex group produced more error-free T-units per T-unit \( (M = 0.36) \) than the simple group \( (M = 0.32) \). The complex group produced less errors per T-unit \( (M = 0.96) \) than the simple group \( (M = 1.10) \). However, these \( t \)-tests show no statistically significant difference between the complex and the simple group in both
measures of accuracy (EFT/T; $t = -1.08, p > .05$; and E/T; $t = 1.53, p > .05$, respectively). In other words, task complexity did not affect accuracy of the participants’ production.

For fluency, as shown in Table 5, the complex group had more fluent production than the simple group, as indicated by both measures of fluency. The complex group produced a greater number of words per T-unit ($M = 8.32$) than the simple group ($M = 7.35$). The difference between the complex group and the simple group was statistically significant ($p = .009$). Also, for total number of words per sentence, the complex group produced more words per sentences ($M = 9.81$) than the simple group ($M = 8.44$). This difference was also statistically significant ($p = .006$).

4.1.2. Discussion

Increasing task complexity along resource-directing dimensions affected neither accuracy nor syntactic complexity of L2 argumentative writings in the present study. These results contradict the prediction based on the Cognition Hypothesis. Robinson (2001a, 2001b, 2007) argues that increasing task complexity along resource-directing dimensions positively affects the quality of L2 performance in accuracy and complexity, because cognitively complex tasks lead learners to focus on linguistic forms and specific linguistic features, making language production more accurate and complex.

If it is assumed that humans have multiple attentional resources to simultaneously process information as the Cognition Hypothesis assumes, the participants who performed the complex task should have had a more accurate and syntactically complex language production, being able to simultaneously process both contents of the task and linguistic features, and conveying meaning without any trade-off effects between complexity and accuracy. However, the results show no effect of task complexity along resource-directing manipulated by +/- no reasoning demands and +/- few elements in accuracy and complexity of L2 argumentative writings in the present study.

Instead, the results are in line with the Limited Attentional Capacity Model, which claims that learners may have limited attentional capacity when they are performing cognitively demanding tasks (Skehan, 1996, 1998; Skehan & Foster, 2001). According to this hypothesis, a complex task overburdens learners with cognitive demands, so they may experience a lack of attentional resources. In this case, they decide how to allocate their limited attention to certain dimension(s) of performance to complete the task. In the present study, the complex group only outperformed the simple group in fluency, but accuracy and syntactic complexity were not affected by the increased task complexity. For example, the complex group produced more words within a T-unit or sentence than simple group by frequently making compound and complex sentences like The first reason for matching them is Minsu and Hyun have introvert personality so that their studying style
won’t make any problem and It’s important because each can help to wake up earlier even one of them is tired. It would appear that the complex task pushed leaners to prioritize fluency, but not accuracy or syntactic complexity during task performance. The participants who completed the complex task primarily allocated their limited attention to fluency in order to properly deliver meaning over linguistic forms. Learners have limited attentional capacities to attend to the dimensions of performance when they are suffering from cognitive demands in complex tasks.

In particular, Skehan (1996) notes that “Tasks which are too difficult are likely to over-emphasize fluency, as learners only have the attentional capacity to convey meanings, using production strategies, lexicalized language, and making meaning primary” (p. 53). In other words, second language learners who are not able to automatically process language tend to concern themselves more with meaning than forms of language in order to deliver what they intend.

William and Graumann (2009) also note that “Lacking sufficient attentional resources to attend to both form and meaning, the latter tends to be prioritized to ensure that the intended message is properly conveyed” (p. 5). That is, concern for meaning rather than language forms is the focus in delivering a message properly when learners have limited attentional capacity. This greater focus on fluency when learners perform cognitively demanding tasks can be also supported by studies claiming that learners’ primary concern is focused on delivering meaning (Skehan & Foster, 2001; VanPatten, 1990). As a result, little attention may be given to linguistic forms as indicated by accuracy and complexity, since learners primarily attend to fluency to convey meaning.

Moreover, the results of the present study are in accord with other empirical studies on the effects of increased task complexity on fluency (Hosseini & Rahimpour, 2010; Ishikawa, 2006; Ong & Zhang, 2010; Salimi et al., 2011). These empirical studies found that complex tasks led to an increase in fluency. Hosseini and Rahimpour (2010) in particular found no effect of task complexity on accuracy and complexity of L2 learners’ written narratives, which mirror the results of this present study. They explain the results according to Skehan and Foster’s (2001) proposition, namely, that prioritization of fluency hinders attention to the other dimensions of accuracy and complexity.

In sum, the greater focus on fluency within the complex group in this present study may be explained by the trade-off effects between fluency and form, as the primary distinction between performance dimensions based on the Limited Attentional Capacity Model, which emphasizes the trade-off effects among different aspects of language performance. That is, the English learners’ attentional resources did not seem to be allocated to linguistic forms, as indicated by accuracy and complexity of performance when the learners were asked to complete the cognitively complex task, since they suffered from a lack of attentional capacities.
4.2. Specific Measures

4.2.1. Results of frequency and target-like use of conjunctions

Frequency, as defined the number of conjunctions per 100 words in this study, is used as a task-specific measure of complexity. Table 6 presents information on the raw number for each conjunction used by the participants and the adjusted number per 100 words. For tokens, the complex group produced a greater number of conjunctions than the simple group. That is, of the 12 selected conjunctions, the complex group produced more conjunctions in their performance than did the simple group for 11 conjunctions, respectively. The greater raw number of conjunctions in the complex group is understandable, because the complex task asked the participants to choose two couples and write at least three reasons for each couple, while the simple task asked the participants to choose one couple. Therefore, in order to ensure comparability, the raw conjunction counts should be converted to conjunction frequencies, which are not biased for text length. In addition, the statistics for frequency of each group require validation.

### TABLE 6

<table>
<thead>
<tr>
<th>Conjunctions</th>
<th>Simple (N = 55)</th>
<th>Complex (N = 55)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw number</td>
<td>Converted number</td>
</tr>
<tr>
<td>and</td>
<td>96</td>
<td>2.16</td>
</tr>
<tr>
<td>so</td>
<td>82</td>
<td>1.85</td>
</tr>
<tr>
<td>because</td>
<td>30</td>
<td>0.70</td>
</tr>
<tr>
<td>but</td>
<td>22</td>
<td>0.52</td>
</tr>
<tr>
<td>if</td>
<td>21</td>
<td>0.49</td>
</tr>
<tr>
<td>when</td>
<td>20</td>
<td>0.48</td>
</tr>
<tr>
<td>so that</td>
<td>2</td>
<td>0.05</td>
</tr>
<tr>
<td>as</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>though</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>although</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>or</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>until</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>273</td>
<td>6.25</td>
</tr>
</tbody>
</table>

Table 7 and Table 8 show the statistics for conjunction frequencies and target-like use of conjunctions, respectively. For frequency, the complex group produced more conjunctions ($M = 6.39$) than the simple group per 100 words ($M = 6.25$). However, no statistical significance was found in a $t$-test ($t = -.26, p > .05$). In other words, the task complexity along the resource-directing dimension did not significantly affect the frequency of each group’s writing performance.
Target-like use of conjunctions is considered a measure of accuracy in the present study. As shown in Table 8, no statistically significant difference between the groups was found in a *t*-test (*t* = -1.86, *p* > 0.05), although the complex group used conjunctions more accurately (*M* = 0.91) than the simple group (*M* = 0.85). As for the frequency data, the variable of target-like conjunctions use did not show any significant effect of task complexity on L2 written performance.

### 4.2.2. Discussion

A few empirical studies have investigated the effects of task complexity on L2 performance by employing task-specific measures. These studies were motivated by the suggestions that task-specific measures would be more sensitive than global measures of CAF to observe the effects of cognitive complexity on performance, and that task-specific measures may be relevant to detect L2 developmental stages for certain linguistic features (Michel, 2013; Revesz, 2011; Robinson, 2005, 2007; Robinson & Gilabert, 2007; William & Graumann, 2009). The present study was also conducted based on the suggestions.

Some studies showed statistically significant effects of task complexity on accuracy and complexity when task-specific measures were used where global measures did not detect the effects of task complexity on L2 performance (Revesz, 2011; Robinson, 2005, 2007; William & Graumann, 2009). On the other hand, Michel (2013) did not find any effect of task complexity on the use of the conjunctions measured by task-specific measures. The present study found task complexity affected neither the frequency as a complexity measure, nor target-like use of conjunctions, as accuracy measures. That is, the complex task did not push learners to accurately or frequently use conjunctions than the simple task.

Although task complexity did not affect the quality of the argumentative writings in frequency and target-like conjunction usage, there is some room for task-specific measures to detect L2 development, since the complex task resulted in more varied use of the
advanced conjunctions than the simple task, as Robinson (2005) would argue. Of course, only 9 percent of the participants were advanced learners in this study, and few advanced conjunctions were used in the complex task. In addition, varied use of the advanced conjunctions may be more relevant to learners’ proficiency levels than the role of the task-specific measure itself. However, it would be worthwhile to conduct empirical research to address this concern. For example, or and until were only present in the complex group, and these are more developmentally advanced conjunctions than and, because, so, but, and when, according to Diessel (2004). He claims that the conjunction and is the first conjunction acquired by children and is followed by other conjunctions such as because, so, but, and when. After that, children acquire coordinating or, conditional if, and the conjunctions marking temporal clauses such as since, after, while, until, and before. However, this suggestion should be taken with careful consideration of the influence of learners’ proficiency levels. Therefore, further studies using task-specific measures are needed to investigate the potential of task-specific measures as detectors for developmental stages for particular linguistic structures.

5. CONCLUSION

The present study aimed to investigate the effects of task complexity on the L2 argumentative writings written by 110 Korean high school students using global measures of CAF and task-specific measures. As a result, the complex task led to more fluent performance than the simple task in the two measures of fluency, using global measures of CAF. It would appear that the learners primarily allocated their attention to fluency rather than to complexity or accuracy in L2 written performance when cognitive demands were imposed upon the learners. Meanwhile, task complexity did not affect the accuracy and syntactic complexity of the written performance. Since the complex task overburdened the learners with cognitive demands, they were not able to allocate additional attention to syntactic complexity and accuracy. For task-specific measures, manipulating task complexity affected neither the frequency nor target-like use of conjunctions.

The results of the present study may suggest some issues for future research. The first would be investigating the relationship between the effects of task complexity on L2 language production and individual difference variables. In addition to task-specific measures, individual variables might influence the effects of task complexity on L2 production (Revesz, 2011; Robinson, 2005).

In addition, the present study used only one type of task-specific measure, conjunctions, and equally weighed each conjunction without examining the specifics of each instance. Thus other types of task-specific measures may also be employed to detect the effects of
task complexity on language performance. Moreover, a closer look at each component of task-specific measures for investigating the effects of task complexity may be warranted. In particular, task-specific measures are considered as developmentally motivated measures (Diessel, 2004; Revesz, 2011). Thus, using task-specific measures based on the developmental stages of language acquisition would also help to detect the effects of task complexity on L2 performance.

Finally, the present study provides some pedagogical implications for second language writing. It was found that learners primarily allocate their limited attention to meaning/fluency instead of linguistic forms when a L2 writing task overburdens learners with cognitive demands. It is hard to improve accuracy or complexity of learners’ output when learners are busy only with delivering meaning. Therefore, instructors need to first think about which dimensions of performance their learners should try to improve and help them achieve a balanced language development by selecting appropriate language learning tasks. It is hoped that the results of the present study and suggestions for future research may contribute to the research on task complexity and task-based language teaching.

REFERENCES


APPENDIX A

Simple Task

기숙사생활을 하는 한 남자고등학교에 신입생이 입학했습니다. 룸메이트를 선정하는데 여러분들의 의견을 반영하려 합니다. 아래에 나와 있는 4명 학생의 특징을 고려하여 최고의 룸메이트가 될 1쌍을 정하세요. 그와 같은 결정을 한 이유를 영어로 보고 하려고 합니다. 본인이 최고의 룸메이트를 고른 이유를 3가지 이상 최대한 자세하게 쓰세요.

최고의 룸메이트

<table>
<thead>
<tr>
<th>민수(Minsu)</th>
<th>지원(Jiwon)</th>
<th>성빈(Sungbin)</th>
<th>도준(Dojun)</th>
</tr>
</thead>
<tbody>
<tr>
<td>취미생활</td>
<td>컴퓨터 게임, 음악 듣기</td>
<td>음악 듣기, SNS</td>
<td>사진 찍기, 요리</td>
</tr>
<tr>
<td>성격</td>
<td>조용하고, 다른 사람의 말을 듣기 좋아함</td>
<td>활발하고, 관람을 잘 하함</td>
<td>다른 사람과 어울려서 노는 것을 좋아함</td>
</tr>
<tr>
<td>취침 시간대</td>
<td>일찍 자고 일찍 일어남</td>
<td>일찍 자고 늦게 일어남</td>
<td>일찍 자고, 늦게 일어남</td>
</tr>
<tr>
<td>공부하는 스타일</td>
<td>혼자 집중하면서 공부함</td>
<td>소리 내서 입으면서 공부하는 스타일</td>
<td>다른 사람들과 같이 공부하는 것을 좋아함</td>
</tr>
</tbody>
</table>
기숙사생활을 하는 한 남자고등학교에 신입생이 입학했습니다. 품매트를 선정하는데 여러분들의 의견을 반영하려 합니다. 아래에 나와 있는 6명 학생의 특징을 고려하여 최고의 품매트가 될 2쌍을 정하십시오. 그와 같은 결정을 한 이유를 영어로 보고 하려고 합니다. 본인이 최고의 품매트를 고른 이유를 글로 각각 3가지 이상 자세하게 쓰세요.

최고의 품매트 1 : __________ & __________
최고의 품매트 2 : __________ & __________

<table>
<thead>
<tr>
<th></th>
<th>민수(Minsu)</th>
<th>지원(Jiwon)</th>
<th>성빈(Sungbin)</th>
<th>도준(Dojun)</th>
<th>한진(Hanjin)</th>
<th>현성(Hyunsung)</th>
</tr>
</thead>
<tbody>
<tr>
<td>좋아하는 과목</td>
<td>영어, 음악</td>
<td>수학, 음악</td>
<td>미술, 체육</td>
<td>컴퓨터, 음악</td>
<td>음악, 체육</td>
<td>수학, 컴퓨터</td>
</tr>
<tr>
<td>취미 생활</td>
<td>컴퓨터 게임, 음악, 듣기</td>
<td>음악 듣기, 창의, 취미</td>
<td>그림 그리기, 사진, 음악, 요리</td>
<td>음악 듣기, 요리, 그림 그리기</td>
<td>컴퓨터 게임, 피아노 연주, 음악 듣기</td>
<td>요리, 마술 배우기, 컴퓨터 게임</td>
</tr>
<tr>
<td>성격</td>
<td>조용하고, 다른 사람의 말을 듣는 것을 좋아함</td>
<td>활동적, 친절하고, 남을 잘 도와줌</td>
<td>활동적, 다른 사람과 어울리고 노는 것을 좋아함</td>
<td>활동적, 다른 사람과 어울리고 노는 것을 좋아함</td>
<td>활동적, 다른 사람과 어울리고 노는 것을 좋아함</td>
<td>활동적, 다른 사람과 어울리고 노는 것을 좋아함</td>
</tr>
<tr>
<td>취침 시간대</td>
<td>일찍 자고, 일찍 일어남</td>
<td>늦게 자고, 일찍 일어남</td>
<td>일찍 자고, 일찍 일어남</td>
<td>늦게 자고, 일찍 일어남</td>
<td>일찍 자고, 일찍 일어남</td>
<td>늦게 자고, 일찍 일어남</td>
</tr>
<tr>
<td>청소</td>
<td>잘 안 함</td>
<td>잘 함</td>
<td>보통</td>
<td>잘 함</td>
<td>잘 안 함</td>
<td>보통</td>
</tr>
<tr>
<td>공부하는 스타일</td>
<td>스타일, 독서</td>
<td>독서, 외국어</td>
<td>학문과 공부하는 스타일</td>
<td>학문과 공부하는 스타일</td>
<td>학문과 공부하는 스타일</td>
<td>학문과 공부하는 스타일</td>
</tr>
</tbody>
</table>

Applicable levels: Secondary

Hyejin Cho
Department of English Language and Literature
Graduate School, Sookmyung Women’s University
97-64 Cheongpa-dong 1-ga, Yongsan-gu
Seoul 140-131, Korea
Phone: 02-710-9321
Email: nina8909@naver.com
Received in March 1, 2015
Reviewed in April 15, 2015
Revised version received in May 15, 2015