Journal of English language
Teaching and Learning
University of Tabriz
Volume 11, Issue 23, (Spring & Summer 2019)

Cognitively complex tasks and individual differences: Two influential factors in Iranian EFL learners' written text quality*

Farahman Farrokhi **
Professor of EFL, University of Tabriz
Simin Sattarpour (Corresponding Author) ***
Assistant Professor at Department of Basic Sciences, Faculty of Paramedicine, Tabriz University of Medical Sciences

Abstract
Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity Model provide the major impetus for this study. The present article reports the findings of a between-subject factorial experimental research study which explored 1) the effects of increased cognitive task complexity, manipulated through the intentional reasoning demands and number of elements on the lexical and syntactic complexity, accuracy, and fluency (CAF) of EFL writers' productions; and 2) the joint effects of cognitive task complexity factor and learners' language learning aptitude (Low vs. High) on the written output. Firstly, we gave Carroll and Sapon's Modern Language Aptitude Test (MLAT) to 226 participants and then did a random stratification of the low- and high-aptitude learners into three groups. The participants received letter writing tasks with different cognitive complexity levels (low, medium, and high). The findings indicated that increasing cognitive task complexity resulted in significantly higher lexical and syntactic complexity and lower fluency, whereas no significant effect was found on writing accuracy. Moreover, the statistical results revealed no significant interaction effect between task complexity factors and learners' language aptitude. With regard to the first objective of the study, the findings supported the predictions of Cognition Hypothesis while it is not the case in relation to the second objective of the study.

Keywords: Cognitive complexity, language learning aptitude, lexical complexity, syntactic complexity, accuracy, fluency

*Received date: 2019/04/14 Accepted date: 2019/04/27
**E-mail: ffarrokhi20@yahoo.co.uk
***E-mail: sattarpours@tbzmed.ac.ir

Introduction

Whereas the effect of task related variables on oral language production has been much reported in the past twenty years, investigating the connection between task-based research and writing skill especially the role of task complexity in writing performance is an insufficiently researched area (Carless, 2012; Johnson, 2017; Porte & Richards, 2012; Samuda & Bygate, 2008). It is important to take into account the significant role of tasks in writing processes since they connect the writer with his writing environment and the intended audience, and then bring the value of both real-world application and attention to meaning, content, and linguistic elements to the forefront (Byrnes & Mancho´n, 2014; Ruiz-Funes, 2015). The existing writing models (e.g., Flower & Hayes, 1980; Kellogg, 1996) have proposed that writing is a kind of problem solving activity including some cognitive processes requiring a writer's constant management of limited attentional resources. In this regard, task-based research can be a good way leading us to grasp what learners allocate their limited attentional and memory resources to, how task manipulation affects the cognitive processing of writers, and how this manner of resource allocation may impact the resulting writing quality.

Among the small number of research studies (e.g., Ishikawa, 2006; Kormos, 2011; Kuiken &Vedder, 2007, 2008, 2012; Ong, 2014; Ong & Zhang, 2010, 2013; Shajeri & Izadpanah, 2016; Tavakoli, 2014) in the field of writing task complexity, cognitive burden has been manipulated by different means such as planning time, writing assistance, draft availability, removal of narrative context, here-and-now variables, and reasoning demand. Concerning the abovementioned variation in the manipulated variables, more similar or even partially replicated research studies are needed in order to reach more vigorous findings in the literature (Porte & Richards, 2012).

Statement of the problem

Writing needs to be investigated more by researchers since it is a complicated, meaning- making, problem-solving, and cognitive

activity in which multiple factors such as the learner, teacher, task, resources availability are all at play. Due to the higher possibility for online planning and its long-lasting and unhurried state, writing is believed to permit the writers to maintain their focus on linguistic aspects of production. Therefore, the effects of cognitive task complexity manipulation are likely to be more evident in written modality (Byrnes & Mancho´n, 2014). Accordingly, doing more task-based research through examining the effects of manipulating cognitive task complexity factors can fill the gap in the literature and pave the way for understanding what learners allocate their attentional resources to and how this manner may impact the writing and consequently language learning and development.

Based on the above considerations concerning disconnected findings due to the diversity of the manipulated factors, one of the objectives of the current study was to contribute further empirical data to existing literature through investigating the interplay between foreign language writing and task-based language teaching by manipulating writing task complexity and then observing the resulting differences. This study is a partial replication of Kuiken and Vedder's (2007, 2008, 2012) and Frear and Bitchener's (2015) studies using similar cognitive complexity factors (intentional reasoning demands and number of elements). Furthermore, due to scarcity of research studies investigating the role of IDs along with task complexity factors, the second contribution as well as novelty of the present study lie on examining how language learning aptitude can interact with task complexity factors to differentiate the ways learners' attentional resources are allocated. As asserted by Kormos and Trebits (2012), IDs such as anxiety, working memory capacity and aptitude might affect task performance by having an impact on students' decisions regarding the allocation of their attentional resources, their ability to handle attentional limitations, and also on the way they acquire particular aspects of linguistic competence.

Review of the related literature

First, the concept of task complexity and its differential effects on linguistic elements of the learners' productions is elaborated from the dissimilar perspectives of Robinson (2001a, 2001b, 2003) and Skehan (1998, 2001, 2005) in this section. Second, some recent task complexity research studies more related to our work are presented. Finally, the role of IDs in task-based research studies is explained by referring to the few empirical studies existingf in this field; moreover, the conceptualizations and operationalization of language learning aptitude as an independent factor in this study are briefly clarified.

Cognitive task complexity

The link between tasks' cognitive demands and quality of task performance is one of the notable issues in task-based research. In Robinson's (2001b) definition, "task complexity is the result of attentional, memory, and other information processing demands imposed by the structure of the task on the language learner" (p. 29). In the SLA literature, there are two major conflicting claims regarding how cognitive variables affect students' performance and also how students use their different attentional resources in task completion: Skehan's (1996, 1998, 2001, 2003) Limited Capacity Hypothesis and Robinson's (2001a, 2001b, 2003, 2005b, 2007b, 2011) Cognition Hypothesis.

The Limited Attentional Capacity Model assumed that our information processing capacity is limited, and as cognitive burdens of a task get heavier, more attentional resources are required to complete the task, hence resulting in trade-off effects among the three aspects of language production: accuracy, fluency, and complexity (Skehan & Foster, 1999, 2001, 2005). It is claimed that cognitively complex tasks distract learners' attention from linguistic forms since more attention will be paid to the content of the message, and not all aspects of language performance can be attended at the same time (Skehan, 2009).

In contrast, the Cognition Hypothesis introduced a multiple-resources view of processing in the mind and claimed that learners have capacity to attend to more than one aspect of language at the

same time. In this model, more cognitively complex tasks are assumed to cause a better performance especially with respect to accuracy and complexity. Two main dimensions of task complexity in Robinson's (2001a) Triadic Componential Framework which determine the cognitive load of the tasks are a) 'resource directing' feature which refers to reasoning requirements of a task (e.g., ± reasoning, ± here-and-Now) and b) 'resource depleting' feature which refers to the needs of strategic planning to do a task (± planning time, ± prior knowledge). Robinson (2001a, 2005b, 2007b) claimed that increasing task complexity along resource-directing dimensions will result in more accurate and complex production as learners have to attend to the conceptual or functional demands of the task, but it will lead to a lower fluency, as learners have to purposefully process language. On the other hand, it is postulated that an increase in task complexity with respect to the resource-dispersing dimensions will decrease fluency, accuracy, and complexity because this factor depletes learners' attentional resources and draw their attention away from any particular aspects of the linguistic system. In order to increase written task complexity in the current study, the researchers manipulated intentional reasoning demands and the number of elements of the tasks which are considered as resource-directing factors in the abovementioned framework.

Task complexity studies on written production

To have a more focused review, only research studies exploring the effects of resource-directing factors are presented here. Kuiken and Vedder (2006, 2007, 2008, 2011), in a series of studies, analyzed the impact of manipulating letter-writing task complexity with respect to the number of elements and reasoning demands on syntactic complexity, lexical complexity, and accuracy. The findings showed that increasing task complexity led learners to write more accurately but not syntactically more complex texts. Regarding lexical complexity, Kuiken and Vedder (2006, 2007) found that increasing task complexity enhanced the lexical variation while no significant effects on lexical variation was found by Kuiken and Vedder (2008,

2011). The differences in lexical complexity were only observable in the measures that did not account for text length. Recently, Kuiken and Vedder (2012) also explored the effect of L2 proficiency on the relationship between task complexity and linguistic production based on data from the three previous studies (Kuiken & Vedder, 2007, 2008, 2011). The findings showed that proficiency levels did not have any interaction effects with task complexity on linguistic performance. Following Kuiken and Vedder's studies, Frear and Bitchener (2015) assessed the effects of the same cognitive task complexity on lexical and syntactic complexity. They noticed that though lexical complexity increased as a result of the increase in cognitive task complexity, this was not the case for syntactic complexity. However, some variations were seen in syntactic complexity when the ratios of different dependent clauses to T-units were measured separately.

Using the same resource-directing factors as the above studies, Cho (2015) examined 110 Korean EFL learners' performance on argumentative writing essays. Unlike the predictions of the Cognition Hypothesis, it was found that the participants performing the complex task produced more fluent writings than those doing the simple task, and no significant effect was observed on accuracy or syntactic complexity of the argumentative writings. Finally, Shajeri and Izadpanah (2016) explored the impact of the cognitive complexity manipulation of the narrative picture sequence tasks through reasoning demands on 48 EFL writers' performance. More complex task was found to facilitate the lexical complexity, syntactic complexity, and fluency of the outputs whereas less accuracy was reported in the more complex task. Thus, the findings of this study seem to confirm the existence of a trade-off between accuracy and complexity as predicted in the Limited Attentional Complexity Model.

Investigating different resource-directing factors, Ishikawa (2006) examined the effects of the ± Here-and-Now dimension and found that increasing task complexity led to greater accuracy, complexity, and fluency of written language. More recently, Ong and Zhang (2010) studied the effects of ± draft-availability factor as a resource-directing

factor along with resource-dispersing factors of ± planning time and ± writing assistance on the fluency and lexical complexity in EFL learners' argumentative essays. Regarding ±draft availability, no significant differences in fluency and lexical complexity were discovered. In addition, Ong and Zhang (2013), utilizing the same data, analyzed the effects of the same variables on the writing quality and found that the learners revealed a tendency toward a high quality production in a condition that the first draft was not available to them.

Mohammadzadeh, Dabaghi, Tavakoli (2013) also investigated the effects of increasing task complexity along both Here-and-Now and planning time dimension on fluency, accuracy, and complexity of 30 EFL learners' written performance on narrative picture sequence tasks. Regarding the resource-directing dimension (Here-and-Now), they did not detect any significant effects on any of the target linguistic measures.More recently, Tavakoli (2014) also used two narrative tasks with different levels of CTC and examine their effects on L2 learners' written performance. In less complex task, the story was based on a single storyline including the foreground events while in more complex task, the storyline was unfolded through the events in both the background and foreground of the story. Syntactic complexity was the only measured linguistic element which was calculated through Mean length of T-units. In line with predictions of Cognition Hypothesis, more complex task encouraged more subordination and longer grammatical units.

Furthermore, Ruiz-Funes (2015) attempted to increase writing tasks' complexity through different essay topics and genres which were assumed to impose different degrees of reasoning demands on the writers. She analyzed the data from two separate studies conducted by the same researcher (2013, 2014) at two distinct language proficiency levels: advanced and intermediate. In both studies, syntactic complexity increased while accuracy and fluency decreased on more complex task. Findings also suggested that the relationship between task complexity and linguistic measures may be associated with the language proficiency level in terms of writing

expertise in the way that high-proficient writers benefit more from the cognitively demanding tasks in comparison to low-proficient ones.

It can be inferred from these findings that there is a lack of consistency in the results, and this makes it challenging to make robust claims. The factors contributing to these inconsistent results can be the wide variation of independent factors used to manipulated task complexity and also different measures used in accounting for linguistic elements of language production. Therefore, more similar studies are required in order to consolidate the results.

Individual differences (IDs) and task performance

Research on interactions between IDs and external variables allow for more "fine-tuned predictions that help with adaptation of teaching methodologies to students or matching students with treatments" (DeKeyser, 2012, p.190). Despite a great emphasis on the need to examine these probable interactions, only few studies (Kormos & Trebits, 2011, 2012; Niwa, 2000; Norris & Ortega, 2009; Robinson, 2007a) have addressed the issue of how IDs may differentiate the ways in which learners can benefit from the manipulation of certain task features. Thus, it is indicative of a real scarcity in interaction research studies in SLA field.

Niwa (2000) investigated the relationship between reasoning demand feature of the task complexity and three ID variables: intelligence, aptitude, and working memory. In this study, by increasing the task complexity, a greater differentiating role of cognitive IDs was found in task performance. Robinson (2007b) also analyzed how input, processing, and output anxiety affected students' performance on narrative tasks with different cognitive loads. He concluded that the negative correlation between output anxiety and syntactic complexity got stronger in more complex tasks; however, no significant link between anxiety and accuracy and fluency of task performance was found.

Moreover, the effects of working memory capacity on the narrative tasks' performance were explored by Kormos and Trebits

(2011). Their findings suggested that high working memory capacity might allow students to produce narratives with high clausal complexity, but it might not be conducive to directing learners' attention to specific dimensions of the task such as subordination. In general, it was concluded that this ID variable seems to play a limited role in influencing the learners' output. Finally, Kormos and Trebit (2012) tried to find out the relationship between components of aptitude and the fluency, lexical variety, syntactic complexity, and accuracy in both written and spoken narrative tasks. The results illustrated a complex interaction between aptitude components and task performance under different conditions, and deductive ability and grammatical sensitivity were the components of aptitude strongly relating to the accuracy and complexity of production.

As language learning aptitude has been considered to be one of the best predictors of language learning success (see Abrahamsson & Hyltenstam, 2008; Dornyei, 2005), the researchers of this study intended to examine the interaction effect of language learning aptitude with task complexity features, and the concept of aptitude will be presented more extensively in the following section.

Language learning aptitude

Carroll (1993) referred to aptitude as "a cognitive ability that is possibly predictive of certain kinds of future learning success" (p. 16); however, he had already clarified that aptitude should be defined only in terms of predicted learning rate not the final achievement. Carroll (1981) determined four factors underlying foreign language aptitude:

> i) phonetic coding ability, that is, the ability to identify distinct sounds, to form associations between those sounds and symbols representing them; ii) grammatical sensitivity or the ability to recognize the grammatical functions of words; iii) rote learning ability, defined as the ability to learn associations between sounds and meanings rapidly and efficiently; and iv) deductive learning ability, which is the ability to infer or induce the rules governing a set of language materials, given sample language materials. (p. 105)

In another conceptualization of aptitude, suggested by Robinson (2002, 2007a), language learning aptitude consists of a network of aptitude complexes, and there are interactions between aptitudinal complexes and different contexts for learning. He has identified a number of 'aptitude-complexes' or combinations of cognitive abilities that are differentially related to processing under different conditions of instructional exposure to L2 input, and the strengths in one or another of these complexes of abilities can be expected to be important to learning from one instructional technique, or under one condition, versus another. Although Robinson (2005a) has obviously pointed to a clear need to update the current theories and measures of aptitude, the existing aptitude tests have not followed the assumptions of these new frameworks. Therefore, most of the researchers rely on the most valid existing psychometric aptitude tests.

The Current Study

The purpose of this investigation is twofold: First, to examine the differential effects of increasing cognitive task complexity on lexical and syntactic complexity, accuracy, and fluency of the learners' letter writing tasks and second, to find out about the interaction effects of cognitive task complexity factor and language learning aptitude on the linguistic elements of writing quality. The following research questions are sought in this study:

1. What are the effects of increased task complexity, manipulated through intentional reasoning demands and the number of elements, on the lexical and syntactic complexity, accuracy, and fluency of EFL learners' letter writing task output?
2. What are the joint effects of increased task complexity and learners' language learning aptitude on the lexical and syntactic complexity, accuracy, and fluency of EFL learners' writing task output?

Method

Participants and context

The population from which the sample was obtained consisted of non-native learners of English as a foreign language. The data of the

current study was collected from three state universities in Iran. The participants' age ranged from 19 to 29 years old (*M*=22.64). They (N= 257) were undergraduate students majoring in English literature and English language teaching, and data collection in this study was conducted during normal university class sessions. In order to assess the language proficiency of the participants and check their homogeneity, Oxford Quick Placement Test (OQPT) was given to them before starting to collect the data. Three participants withdrew halfway, and 28 ones, being determined as either highly advanced or beginners through the OQPT scores, were excluded from the study to save the homogeneity of the sampling group. Thus, the main sample included 226 (females=128, males=98) learners selected from intermediate levels.

Instruments

Oxford Quick Placement Test (version 1). This test includes 60 multiple choice items with 30-minute allocated time, and test takers can be placed at different proficiency levels (A1= breakthrough to C2 = mastery) according to the score range.

Modern Language Aptitude Test. To measure language learning aptitude, several instruments already exist; however, to date, Carroll & Sapon's (1959) MLAT is the earliest and most used and validated test based on the four main facets of language aptitude mentioned above. The MLAT has been used in numerous studies since 1950s, and there is a growing consensus of its validity when used as a test of language aptitude (see Carroll, 1981, 1993; Como et al., 2002; Parry & Child, 1990; Stansfield & Reed, 2004). DeKeyser (2000) clearly stated that the MLAT "is usually considered the best verbal aptitude test in terms of its predictive validity for L2 learning" (p. 509). Therefore, the MLAT provides a strong measure of language aptitude and is used in the current research study.

In spite of the fact that the MLAT is considered by some to be outdated; it has a very high correlation to general intelligence and is cited as a reliable measure across studies, in various languages, and in different contexts. There is a general consensus among researchers

that there is no clear evidence in the existing literature to prove the superiority of the new testing instruments over the MLAT (Sawyer & Ranta, 2001; Sparks & Ganschow, 2001). Furthermore, some test developers, themselves, admitted the defeat of their tests. For instance, Peterson and Al-Haik (1976) administered the Defense Language Aptitude Battery (DLAB) which was suitable for discriminating better among higher aptitude students, to over 1,000 subjects to compensate for the plateau effect of MLAT (Kiss & Nikolov, 2005, as cited in Dornyei, 2005), but it was found that the predictive validity of the instrument increased not significantly. In a study comparing various aptitude batteries to their newly developed VORD, Parry and Child (1990) also found that the MLAT was "the best overall instrument for predicting language-learning success" (p. 52).

Thus, the MLAT provides a strong measure of language aptitude and will be used in the current research study. Anderson (2012) has summarized four research-based merits of the MLAT as the following:

> 1) the MLAT does in fact measure a different construct than general intelligence; 2) language aptitude, as measured by the MLAT, is not trainable; 3) the MLAT successfully measures the language aptitude of learners in both formal and informal contexts; 4) for 50 years the MLAT has proven to be a valid and reliable measure with 0.4 – 0.6 validity and .83 to .93reliability scores reported (Carroll & Sapon, 2002). Overall, the MLAT continues to be one of the best predictors for success in language learning available today and as such, is an important tool for researchers in SLA and the fields of language education and educational psychology (p. 57).

The MLAT is comprised of five subtests. Part 1, Number Learning, which targets timed associative learning, asks participants to learn the names of 1-, 2-, and 3-digit numbers in a new language and then transcribe the numbers they hear. Part 2, Phonetic Script, tests participants' phonemic coding ability. Test takers study a phonetic script and choose the word that they hear from choices written in

phonetic script. Part 3, Spelling Cues, taps participants' vocabulary knowledge and their ability to handle novel spellings of known words. Part 4, Words in Sentences, targets language analytic ability and grammatical sensitivity. Test takers are required to choose a word that serves the same grammatical function as a specified word. Part 5, Paired Associates, targets rote memory learning. Test takers study a list of Kurdish vocabulary items and their English translations, and then they complete a multiple-choice test of the word pairs without referring to the original lists.

Since the original version of the MLAT including 149 items was very long and tedious, the participants were administered a shortened version of this test. Therefore, five parts with overall 90 items including, number learning (15 items), phonetic script (15 items), spelling clues (20 items), words in sentences (20 items), and paired associates (20 items) were taken from the original long version of MLAT.

Letter Writing Tasks. Three letter writing tasks, Task 1 (low complexity), Task 2 (medium complexity), and Task 3 (high complexity), were used for collecting the written data. These tasks were taken from Frear and Bitchener's (2015) work with the full permission of the authors given through E-mail correspondence. However, some parts of the tasks were slightly modified to make them more culturally suitable and tangible for Iranian participants. For instance, foreign currency (dollar), alcoholic beverages, pork meat, and weekend days (Saturday and Sunday) were replaced by the local currency, nonalcoholic beverages, seafood, and local weekend days respectively.

Task 1 obviously entailed the lowest complexity, and it was done purposefully to make sure that the levels of cognitive demand being applied were low enough in comparison to the other two tasks in order to avoid the probable negative influence of inaccurate measurement of cognitive task complexity on results when compared to more complex tasks. In Task 1, the participants were asked to write to an English-speaking friend traveling to their country. They were required to use

their resources and to write to this friend and tell him about Iran. The task instructions were designed to be easily comprehensible while avoiding phrases that instructed the writer to form any opinions or state any reasons why it would be worth moving to their country. In fact, the participants were expected to rely on their own resources. Through manipulating the reasoning demands and the number of elements in the task instructions, Task 2 was made more complex. This was expected to initiate higher amounts of attention demanding activity (Ellis, 2003; Robinson, 2005, 2007b) in order to complete the task. Like Task 1, the participants were required to write to a friend coming to Iran for the weekend, but they were also asked to inform him that which of the two restaurants they would visit upon his arrival and why. In order to complete the task, the participants were expected to take into account the information about two restaurants and also the preferences of the visiting friend provided in the instructions. The most complex task (Task 3) was similar to Task 2, but this time, the participants were supposed to choose one restaurant out of three restaurants not two, and they were also expected to consider the preferences of two more friends who would be visiting the restaurant as well as the foreign visitor.

Target measures: CAF

In task-based research field, the performance is generally measured in terms of CAF. Following the trend, four target measures used in this study are: lexical complexity, syntactic complexity, accuracy, and fluency. According to Ellis and Barkhuizen (2005), it is essential to consider more than one type of complexity when examining the written complexity since learners may use different ways to express complexity in their performance. Therefore, two types of complexity, lexical and syntactic, were measured in this study. Lexical complexity was measured by a mean segmental type-token ratio which takes variety in the length of texts into account; and syntactic complexity was measured through the ratio of dependent clauses to T-units. T-units are frequently used in the literature and are proposed to be good indicators of the progress in writing ability (Hunt, 1965). Accuracy of

the performance in these writing tasks was measured with the ratio of error-free T-units to the total number of T-units. Any violations of syntactical, morphological, and lexical norms, but not punctuation and capitalization ones, were considered as error. Finally, fluency was measured by calculating the mean number of words produced per minute out of the total time the participants spent on the task.

Design of the study

In order to collect the quantitative data of the present study, one independent variable (intentional reasoning demands) was manipulated and one more independent variable (language leaning aptitude) was measured. The design of this between-subjects study was experimental factorial one since there were more than one independent variable. The purpose of this design is to study the independent and also interaction effects of two or more independents variables on an outcome (Vogt, 2005). Because three levels of reasoning demands and two levels of language learning aptitude were measured in the study, the design was called 3 ×2 factorial design

Procedures

A week prior to the beginning of the study, Oxford Quick Placement Test was administered to determine the participants' proficiency levels and to select those who could go under investigation. In the first data collection session, the participants took the MLAT which lasted about 30 minutes along with the researchers' instructions and clarifications. Since MLAT is a speed test, the Participants took all five sections of the MLAT in specified allocated time kept by both proctor and recorded audio. Aptitude score was total number of correct answers in the entire test, and then those gaining the score above the median were identified as high aptitude learner and the others scoring below the median were recognized as low aptitude ones.

Afterwards, a random stratification of the participants was done regarding their aptitude levels into three cognitive task complexity groups. In other words, the number of the participants from each

aptitude level was almost the same in all three experimental conditions. In the last data collection session, Task 1 was given to the first experimental group, Task 2 to the second experimental group, and Task 3 to the third group. After giving the general instructions, the participants were provided with 2 minutes to view the task and ask questions about any aspects of the instructions, and 30 minutes to complete the task. Since lexical complexity was one of the target measures in the current study, no dictionaries or smart phones were allowed.

Data analyses

In order to measure the target elements of written production, all the letters were first typed in MS word documents and coded according to the measures for lexical and syntactic complexity, accuracy, and fluency as mentioned above. The following values were counted and computed for each written production: (a) total number of words (b) total number of T-units; (c) total number of dependent clauses (d) ratio of dependent clauses per T-unit; (e) total number of error-free T-units.

For the analysis of lexical variation, each text was divided into segments of 40 words to avoid wasting data as much as possible. Type-token ratios for these segments were calculated by Text Inspector, a professional web tool for analyzing texts, and then the mean of these ratios were calculated for each text. To measure syntactic complexity, the total number of dependent clauses in a text was divided by the total number of the T-units. Furthermore, accuracy measure, ratio of error-free T-units to the number T-units, was calculated by counting the number of T-units without any errors and dividing it to the total number of T-units in a text. Finally, fluency, for each text, was calculated by dividing the total number of words counted by the same above-mentioned online tool to 30 minutes. About one third of the letters were also coded by a trained PhD candidate in English Language Teaching, and the inter-rater reliability coefficients of .88 and .90 were achieved for syntactic complexity and accuracy. The same was not done for lexical complexity and fluency

since they were measured by the use of a computer tool, so were not subject to any measurement discrepancies.

Results

Our first research question sought the effects of increased task complexity, manipulated through the intentional reasoning demands and number of elements on the lexical and syntactic complexity, accuracy, and fluency of the participants' written outputs. First of all, the descriptive statistics including the mean and standard deviation of the measures of lexical complexity, syntactic complexity, accuracy, and fluency in three different task complexity conditions are shown in Table 1.

Table 1
Means and standard deviations of linguistic measures in three tasks

	Low Complexity			Medium Complexity			High Complexity	
	N	M	SD	N	M	SD	N	M SD
Lexical Complexity	70	.79	.03	76	.77	.03	80	.81 .05
Syntactic Complexity	70	.40	.15	76	.47	.15	80	.41 .17
Accuracy	70	.59	.15	76	.62	.19	80	.61 .16
Fluency	70	8.05	1.62	76	7.01	1.85	80	6.87 2

In order to determine if the participants performed differently under different cognitive complexity conditions, a one-way between-subject MANOVA was conducted on the linguistic target measures of the study. The results revealed that the task conditions significantly affected the fluency,[F(2, 223)=9.25, p = .000, η_p^2=.098], lexical complexity,[F(2, 223)=3.99, p = .020, η_p^2=.046], and syntactic complexity[F(2, 223)=4.57, p = .011, η_p^2=.051], whereas no significant mean scores variation was found regarding accuracy [F(2, 223)=.383, p = .682, η_p^2=.004]. Post hoc pair-wise comparisons using Bonferroni test at an alpha level of .05 were also performed. For fluency, it indicated that the mean score for low complexity condition

(M=8.05, SD= 1.62) was significantly higher than medium complexity condition (M= 7.01, SD=1.85), *p*=.003 as well as high complexity condition (M=6.87, SD=2), *p*=.001. However, the pairwise comparison of the medium complexity condition with high complexity one was non-significant.

Comparing the lexical complexity mean scores, post hoc analyses showed that the writers in the high complexity condition (M = .81, SD = .05) reported a marginally significantly higher mean than the writers in the medium complexity condition (M = .77, SD = .03), p = .028, but the writers in other groups, compared pairwise, didn't show significantly different performances. Finally, regarding syntactic complexity, the mean score of medium complexity condition (M=.47, SD=.15) was found to be significantly higher than the mean score of low complexity condition (M=.40, SD=.15), p=.040. No other mean difference was significant.

To answer the second research question concerning if cognitive task complexity factor and language aptitude interact to influence linguistic elements of the participants' written outputs including lexical and syntactic complexity, accuracy, and fluency, a two-way MANOVA test was conducted. Although a main effects of task complexity conditions[Wilks's Lambda=.85, F(8, 422)=4.47,p=.000, η_p^2=.078] and language aptitude [Wilks's Lambda=.72, F(4, 211)=19.76,p=.000, η_p^2=.273 were found, no interaction effect of cognitive task complexity and language aptitude was observed [Wilks's Lambda=.974, F(8, 422)=.70, p = .69, η_p^2=.013]. Table 2 shows the main and interaction effects of the independent variables on target measure separately.

Table 2

Main and interaction effects of factor one and factor two on linguistic measures

Dependent variables	Task complexity				Aptitude				Interaction			
	F	df	p	η^2	F	df	p	η^2	F	df	p	η^2

Lexical Complexity	3.99	2	.020	.046	19.1	1	.000	.082	.271	2	.763	.003
Syntactic Complexity	4.57	2	.011	.051	4.64	1	.032	0.21	.710	2	.493	.007
Accuracy	.383	2	.682	.004	58.1	1	.000	.214	.260	2	.771	.002
Fluency	9.25	2	.000	.098	.579	1	.448	.003	1.38	2	.253	.013

Since the effects of factor one (task complexity conditions) on linguistic elements are discussed above, we move to clarifying the effects of language aptitude and interaction effects of these two factors. As it can be seen in Table 2, aptitude significantly affected the lexical complexity, syntactic complexity, and accuracy but not fluency. Regarding interaction effects, no significant effect was found on the target measures of the study. Therefore, it can be concluded that learners' language aptitude level didn't mediate in the effects cognitive task complexity on their written performance.

Discussion

The first objective of the study was to explore the effects of manipulating cognitive task complexity with respect to the resource-directing dimensions, reasoning demands and the number of elements. Our results are more seemingly in line with the predictions of the Cognition Hypothesis. Taken together, the present study revealed that increasing task complexity led to some positive effects on lexical and syntactic complexity, negative effects on fluency, but no effects on accuracy. Although accuracy was found not to change significantly across the groups doing the tasks with different complexity levels, the observed mean values showed a tendency of higher values of accuracy in more complex tasks. Therefore, it would be premature to reject Robinson's hypothesis.

As mentioned above, the findings support the predictions of Cognition Hypothesis with respect to lexical complexity. Robinson (2007b) argued that increased task complexity through resource-

directing factors will lead the learners to use a greater range of morphology and lexicon in order to meet the conceptual requirements of the tasks, and as a result it will enlarge their interlanguage system. This phenomenon can be also associated with Schmidt's (2001) justification that expending the mental effort in more cognitively complex tasks prime L2 learners and direct their attentional and memory resources, not only to syntactic but also to lexical and morphological aspects of the L2 system, thereby facilitating selective attention and then noticing.

Focusing on the studies manipulating resource-directing dimensions, the results are in line with the findings of Frear and Bitchener (2015), Ishikawa (2006), Kuiken and Vedder (2006, 2007), and Shajeri and Izadpanah's (2016) studies in which higher lexical complexity measures were found in more complex tasks. On the other hand, unlike the results of this study, Kormos (2011), Kuiken and Vedder (2008, 2011), and Ong and Zhang (2010) did not observe statistically meaningful effects on the lexical complexity as a result of manipulating resource-directing factors.

The contrasting results may be due to the diversity of the complexity factors and also target measures used in the above studies. For instance, although the complexity dimension used by Kormos (2011) was categorized as ± reasoning demands factor in Johnson's (2017) synthesis analysis, we believe that the tasks were completely different from the present study's, and the causal and spatial reason were involved in the tasks rather than intentional reasoning. Furthermore, the letter writing tasks utilized in Kuiken and Vedder's (2006, 2007, 2008, 2011) series were similar to the tasks in this study, but due to the absence of low complexity task (i.e. Task 1 in the present study) in their study, the only cognitively distinguishing factor between their tasks was the number of the conditions not reasoning. Johnson also categorized their studies as the ones utilizing ± few elements feature of CTC. Finally, concerned with Ong and Zhang's (2010) study, draft availability, for the first time, was considered as a resource-directing factor by the authors.

What is also interesting about the results of the current study is that lexical complexity was recognized to be significantly different between only medium and high complexity groups, and the participants in low and high complexity conditions, with the sharpest complexity contrast, did not perform significantly different. Non-significant difference between the lowest and highest complex tasks in lexical complexity can be justified based on the requirements of Task 1. Although high cognitive demands of Task 3 led to more lexical variation, the writers of Task 1 were also asked to freely write about their country from different points of view as they wish which might have led them to utilize a various range of vocabulary. Additionally, these between-group differences found in the present study could be taken as a good reason to take issue with Frear and Bitchener (2015). They suspected that gaining higher lexical complexity in more complex tasks might be due to the higher variation of lexical items in more complex tasks' instructions not merely because of increased CTC effects. In this study, significant differences were detected only between Task 2 and 3, whereas the widest contrast between the number of lexical items in task instructions was obviously between Task 1 and 3. Therefore, we believe that the role of increased cognitive loads of the tasks was more prominent in directing the writers' attention to their lexical repertoire and retrieving them rather than the increased number of given lexical items in instructions.

Concerning syntactic complexity, higher syntactic complexity levels were detected in the more cognitively demanding tasks in corroboration with the predictions of Cognition Hypothesis. The results of this study are similar to those of Ishikawa (2006), Ruiz-Fune (2015), Shajeri and Izadpanah (2016) in terms of obtaining positive effects for increased task complexity. There are also some studies (e.g., Frear and Bitchener, 2015; Kuiken and Vedder, 2006, 2007b, 2008, 2011) in which syntactic complexity did not differ significantly across tasks with different levels of cognitive complexity. The justification behind the obtained discrepancies might be different contexts in which the studies were done, small sample size in their

studies, different research designs (between-subjects and within-subjects designs), or different second languages being learned by the participant for instance Italian and French in Kuiken and Vedder's studies.

Interestingly, unlike the between-group differences in lexical complexity, the participants in both Task 2 and Task 3 conditions outperformed the ones in Task 1 condition involving the lowest cognitive demands. The justification of this significant difference in syntactic complexity between complex tasks (Tasks 2 and 3) and Task 1 can be the different requirements of Task 1 in which the writers only described individual issues in isolation. Thus, low complexity task instructions avoided utilizing words or phrases that might elicit any opinions or reasoning demands expressed by using more subordinations while interrelated issues in Task 2 and 3 accompanied the use of embedded and subordinating means mostly in form of expressing cause and effect structures to support their choice of the best restaurant to go.

With respect to accuracy, no significant difference was observed among three complexity conditions. Nonetheless, we cannot claim that this result totally rejects Robinson's prediction and supports Skehan's trade-off hypothesis since no negative effect was found either, and the observed mean values showed a tendency of higher values of accuracy on more complex tasks. Unlike our study, Ishikawa (2006) and Kuiken and Vedder (2006, 2007, 2008, 2011) determined increased cognitive task complexity as a help to learners for producing more accurate written output but not more complex one. In contrast, Ruiz-Fune (2013, 2014, 2015) found a negative relationship between task complexity and written accuracy. She argued that the attention to language complexity, which was boosted through cognitive task complexity increase, could be considered as a restraint on focusing on accuracy demands simultaneously. However, we hold the idea that injecting some other factors such as planning time or even increasing cognitive demand's levels on the tasks could lead to more control over already existing interlanguage system to write more accurately.

Concerning fluency, the predictions of both Cognition Hypothesis and 'trade-off' Hypothesis were met since moving from low to medium and then to high complex task, the participants wrote less fluently. Similarly, Ong and Zhang (2010) and Ruiz-Funes (2013, 2014, 2015) found that fluency measure appeared to decrease by increasing CTC through reasoning demands. Ruiz-Funes argued that attention to syntactic complexity prevented the participants from simultaneously focusing on and fluency demands. Under Robinson's (2007b) view, fluency requires learners to draw on their memory-based system, accessing and using ready-made chunks of language, thus while their attention is directed to more complex linguistic production which consequently leads to more grammaticisation, attention to higher speed of speaking or writing will be deteriorated.

Concerning the second objective of the study, the Cognition Hypothesis predicts that IDs will increasingly affect task-based performance and learning as tasks increase in complexity, and particularly the learners high in the abilities that resource-directing variables such as reasoning draw on may benefit more (Niwa, 2000; Robinson, 2001a, b, 2005a, b). Although increased task complexity factor seemed to play a significantly effective role in writing quality, no interaction between language aptitude and increased task complexity factor was observed. Thus, based on our results, manipulating task features has similar effects on all learners' cognitive processes in dealing with the task, no matter which level of language aptitude -Low vs. High- they enjoy.

Kormos and Trebits (2011) examining the effects of working memory capacity on written narrative tasks also concluded that ID variable seems not to play a strong role in influencing the learners' output, but it might only lead to narratives with high clausal complexity. Kormos and Trebit (2012) also explored the relationship between components of aptitude and linguistic elements of learners' performance in both written and spoken mode which led to the discovery that two components of aptitude strongly related to the accuracy and complexity of production were deductive ability and

grammatical sensitivity. Generally, it can be inferred that the potential of writing to assist the learners to learn a language is more related to the degree of cognitive complexity involved in writing tasks rather than individual differences, and this might be related to the problem-solving nature of the tasks prompting the learners to use their own cognitive processes in different ways. Another explanation for this obtained result might be related to not distinguishing different components of the aptitude test, aptitude complexes as introduced by Robinson (2002, 2007a), and considering the total score of different components as a unitary signification of the language learning aptitude.

Conclusion

To sum up, the main findings of our study are listed here. Firstly, it was found that complex tasks resulted in significantly greater lexical and syntactic complexity in comparison to less complex tasks which confirms Robinson's predictions in Cognition Hypothesis. Secondly, fluency significantly decreased due to the increased cognitive load of tasks, and this result strongly supports the predictions of both Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity. Thirdly, with respect to accuracy, the prediction of Cognition Hypothesis was supported although not strongly enough to be statistically significant. However, this cannot be regarded as a support of Skehan's hypothesis because negative effect of complex tasks on accuracy was not observed. Finally, no interaction effect between cognitive task complexity and language aptitude was found, which refers to more prominent role of external cognitive factors rather than internal ones.

Pedagogical Implications

The current study provides some pedagogical implications for second language acquisition and second language writing researchers, teachers, syllabus and task designers, and language testing specialists. It was found that when the writers were burdened with more cognitive demands, they primarily allocate their attention to linguistic forms especially complexity and hardly attend to fluent delivering of the

message. Thus, teachers and syllabus designers should consider which linguistic dimensions their learners need to improve first, and if there is no priority to improve specific linguistic dimensions. The teachers should also try to achieve a balanced language development by selecting appropriate language learning tasks. Moreover, it seems that taking into account the level of learners' language learning abilities besides task complexity factors would not lead to a much more productive linguistic result. However, it is not very logical to suggest the teachers to overlook the cognitive capabilities of the learners since IDs, especially cognitive ones, have always been found to impose much more beneficial effects on the learners' performance. We recommend researchers to consider other cognitive IDs in their task-based studies.

The cognitive complexity of a task should also be considered by the testers when they design tasks for testing purposes since they had better know which aspects of the test results are more likely to be affected by task factors. Furthermore, the current study can help task-based researchers to gain more consistent results in writing field which has always been much less attended in comparison to speaking field.

Limitations and Future Research

Despite the potential contributions of the current study, it also suffers from some limitations like any other research study. First, between-subject design of the study might be regarded as a limitation though we tried to compensate for this limitation by giving the participants a placement test and doing a stratified sampling to make sure of their homogeneity. Second, utilizing single measurement for each linguistic element may be another limitation of our study. Thus, using additional measures needs to be taken into account in the future research on writing complexity since different measures are found to lead to some contrastive results in some existing studies in the literature. Third, the individuals' aptitude level was determined by their overall score on the MLAT test which might ignore the various roles of aptitude components under different task conditions as proposed by

Robinson's (2001a, 2005a) Aptitude Complexes Hypothesis. It is also recommended to include the learner affective factors such as motivation, anxiety, and self-efficacy in similar experiments to explore the three-way interplay among task complexity features, cognitive and affective learner variables. Furthermore, it might be worthwhile to investigate the impact of providing different amounts of planning time as a resource-dispersing factor along with the factors manipulated in this study.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30,* 481-509. doi:10.1017/s027226310808073x

Byrnes, H., & Mancho´n, R. M. (Eds.). (2014). *Task-based language learning: Insights from and for L2 writing.* Philadelphia/Amsterdam: John Benjamins. doi:10.1075/tblt.7.01byr

Carless, D. (2012). TBLT in EFL settings: Looking back and moving forward. In A. Shehadeh & C. A. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 345-358). Amsterdam: John Benjamin.doi: 10.1075/tblt.4.20car

Carroll, J. (1981). Twenty-five years of research on foreign language aptitude. In Karl C. Diller (Ed.), *Individual differences and univerals in language learning aptitude* (pp. 119–154). Rowley, MA: Newbury House.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press. doi: 10.1017/cbo9780511571312

Carroll, J., & Sapon, S. M. (1959). *The Modern Language Aptitude Test*. San Antonio, TX: Psychological Corporation.

Como, L., Cronback, L. J., Kupermintz, H., Lohman, D. F., Mandinach, E. B., Porteus, A. W., & Talbert, J. E. (2002). *Remaking the concept of aptitude: Extending the legacy of Richard E. Snow.* Mahwah, N.J.: Erlbaum.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499-533.

DeKeyser, R. M. (2012). Interactions between Individual Differences, Treatments, and Structures in SLA. *Language Learning*, 62(2), 189-200. doi: 10.1111/j.1467-9922.2012.00712.x

Dornyei, Z. (2005). *The psychology of the language learner*. Mahwah NJ: Lawrence Erlbaum. doi: 10.4324/9781410613349

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, 31, 21–32. doi: 10.2307/356630

Frear, M. W., & Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing, 30,* 45-57. doi: 10.1016/j.jslw.2015.08.009

Hunt, K.W. (1965). *Grammatical structures written at three grade levels.* Urbana, IL: The National Council of Teachers of English.

Ishikawa, T. (2006). The effect of manipulating task complexity along the (_Here-and_Now) dimension on L2 written narrative discourse. In C. M. Garcı´a-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136–156). Clevedon, UK: Multilingual Matters.

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing, 37,* 13-38. doi: 10.1016/j.jslw.2017.06.001

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing,* 20, 148–161. doi: 10.1016/j.jslw.2011.02.001

Kormos, J., & Trebits, A. (2011). Working memory capacity and narrative task performance. In P. Robinson (ed.), Second language task complexity: *Researching the cognition hypothesis of language learning and performance* (pp. 268-85).Amsterdam: John Benjamins. doi: 10.1075/tblt.2.17ch10

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality and aptitude in narrative task performance. *Language Learning,* 62(2), 439-472. doi: 10.1111/j.1467-9922.2012.00695.x

Kuiken, F., & Vedder, I. (2006). Cognitive task complexity and linguistic performance in French L2 writing. In M. Garcı´a-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 117–135). Clevedon, UK: Multilingual Matters.

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching,* 45, 261–284. doi: 10.1515/iral.2007.012

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17,* 48-60. doi: 10.1016/j.jslw.2007.08.003

Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 91–104). Amsterdam: John Benjamins. doi:10.1075/tblt.2.09ch4

Kuiken, F., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 143–170). Amsterdam: John Benjamins. doi:10.1075/lllt.32.07kui

Mohammadzadeh M. A. R., Dabaghi, A., & Tavakoli, M. (2013). The effects of simultaneous use of pre-planning along ±here-and-now dimension on fluency, complexity, and accuracy of Iranian EFL learners' written

performance. *International Journal of Research Studies in Language Learning, 2*(3), 49–65. doi: 10.5861/ijrsll.2012.168

Niwa, Y. (2000). Reasoning demands of L2 tasks and L2 narrative production: Effects of individual differences in working memory, intelligence, and aptitude (Unpublished M.A. dissertation). Aoyama Gakuin University, Tokyo.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 555-578. doi:10.1093/applin/amp044

Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing*, 23, 17–30. doi:10.1016/j.jslw.2013.10.002

Ong, J., &Zhang, L.J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, 19, 218–233. doi: 10.1016/j.jslw.2010.10.003

Ong, J., & Zhang, L. J. (2013). Effects of the manipulation of cognitive processes on EFL writers' text quality. *TESOL Quarterly*, 47, 375–398. doi:10.1002/tesq.55

Parry, T. S., & Child, J. R. (1990). Preliminary investigation of the relationship between *VORD*, *MLAT*, and language proficiency. In T. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 30–66). Englewood Cliffs, NJ: Prentice Hall.

Petersen, C., & Al-Haik, A. (1976). The development of the defense language aptitude battery. *Educational and Psychological Measurement, 36*, 369–380. doi: 10.1177/001316447603600216

Porte, R., & Richards, K. (2012). Focus article: Replication in second language writing research. *Journal of Second Language Research*, 21, 284–293. doi: 10.1016/j.jslw.2012.05.002

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for investigating task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). New York: Cambridge University Press. doi: 10.1017/cbo9781139524780.012

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57. doi:10.1093/applin/22.1.27

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies*, 21, 45–105.

Robinson, P. (2005a). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 45-73. doi:10.1017/s0267190505000036

Robinson, P. (2005b). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics, 43,* 1-32. doi: 10.1515/iral.2005.43.1.1

Robinson, P. (2007a). Aptitudes, abilities, contexts and practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from cognitive psychology and applied linguistics* (pp. 256- 286). Cambridge: Cambridge University Press. doi: 10.1017/cbo9780511667275.015

Robinson, P. (2007b). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics, 45,* 237-257. doi: 10.1515/iral.2007.009

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. 3–38). Amsterdam: John Benjamins. doi: 10.1075/tblt.2.05ch1

Ruiz-Funes, M. (2013, November). The interplay between task complexity in foreign language writing at the intermediate level and measures of syntactic complexity, accuracy and fluency (CAF). Paper presented at the American Council on Teaching of Foreign Languages (ACTFL).

Ruiz-Funes, M. (2014). Task complexity and linguistic performance in advanced college-level foreign language writing. In H. Byrnes & R. M. Mancho´n (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 163–192). Amsterdam: John Benjamins. doi:10.1075/tblt.7.07rui

Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing* , 28 , 1–19. doi:10.1016/j.jslw.2015.02.001

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning.* New York: Palgrave Macmillan. doi:10.1057/9780230596429_5

Sawyer, M. & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319-353). New York: Cambridge University Press. doi: 10.1017/cbo9781139524780.013

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language learning* (pp. 3–32). Cambridge: Cambridge University Press. doi:10.1017/cbo9781139524780.003

Shajeri, E., & Izadpanah, S. (2016). The impact of task complexity along single task dimension on Iranian EFL learners' writing production. *Theory and Practice in Language Studies, 6*, 935–945. doi:10.17507/tpls.0605.04

Skehan, P. (1996). A framework for the implementation of task based instruction. *Applied Linguistics,* 17, 38–62. doi:10.1075/tblt.1.06afr

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

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 167–185). London: Longman.

Skehan, P. (2003). Task-based instruction. *Language Teaching,* 36, 1–14. doi: 10.1017/s026144480200188x

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30,* 510-532. doi: 10.1093/applin/amp047

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning,* 49, 93–120. doi:10.1111/1467-9922.00071

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183–205). New York: Cambridge University Press. doi: 10.1017/cbo9781139524780.009

Skehan, P., &Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 193–218). Amsterdam: John Benjamins. doi:10.1075/lllt.11.12ske

Sparks, R. L., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics, 21*, 90–111. doi: 10.1111/j.1540-4781.1991.tb01076.x

Stansfield, C. W., & Reed, D. J. (2004). The story behind the Modem Language Aptitude Test: An interview with John B. Carroll (1916-2003). *Language Assessment Quarterly, 1*(l), 43-56. doi: 10.1207/s15434311laq0101_4

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes, & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 217–236). Philadelphia/Amsterdam: John Benjamins. doi:10.1075/tblt.7.09tav