Content list available at www.urmia.ac.ir/ijltr

Iranian Journal of Language Teaching Research

Urmia University

The effect of task complexity on the quality of EFL learners' argumentative writing

Karim Sadeghi [a,*], Zahra Mosalli [a]

[a] *Urmia University, Iran*

ABSTRACT

Based on Robinson's (2005) Cognition Hypothesis and Skehan and Foster's (2001) Limited Attentional Capacity Model, the current study attempted to investigate the effect of manipulating task complexity on argumentative writing quality in terms of lexical complexity, fluency, grammatical accuracy, and syntactic complexity. Task complexity was manipulated through applying resource-dispersing dimensions. All 60 participants who were university students were randomly assigned into one of the three groups: (a) topic; (b) topic + idea; and (c) topic + idea + discourse marker group. A series of one-way ANOVAs was utilized to detect significant differences among the groups. Results showed that increasing task complexity: 1. did not lead to differences in lexical complexity (measured by the ratio of lexical words to function words and lexical density), but it did lead to significant differences when mean segmental type-token ratio was used to measure lexical complexity; 2. produced significantly less fluent language; 3. resulted in more grammatically accurate language in the least complex task; and 4. did demonstrate significant difference in syntactic complexity (when it was measured by the ratio of dependent clauses to total clauses). Further findings and implications are discussed in the paper.

Keywords: task complexity; lexical complexity; fluency; grammatical accuracy; syntactic complexity; argumentative writing

ARTICLE SUMMARY

Received: 12 Jan. 2013 **Revised version received:** 24 Apr. 2013

Accepted: 28 Apr. 2013 **Available online:** 15 May 2013

* Corresponding author: Urmia University, Iran
Email address: ksadeghi03@gmail.com

Introduction

Among the four skills, writing is the most difficult for foreign language (FL) learners to learn as it requires paying attention to both higher and lower level skills at the same time during the writing process. One of the test methods for assessing writing performance is a 'task' (Bae & Bachman, 2010) which has been considered as a key and indispensable instructional tool in FL learning classrooms. The paramount importance of task has directed many researchers' attention towards task-based language learning, teaching, and research (e.g. Kuiken & Vedder, 2008; Kormos, 2011; Ong & Zhang, 2010).

Task-based research has concentrated mainly on learners' (mental) involvement in task completion process. "Proposals for task-based approach to pedagogy have conceded that valid criteria for determining the difficulty level of tasks have yet to be established" (Robinson, Chi-chien Ting, & Urwin, 1995, p. 62). Regarding theoretical perspectives, there are different writing models (e.g. Flower & Hayes, 1980; Bereiter & Scardamalia, 1987; and Kellogg 1996), but none of these models predicts the nature of processes involved in learners' mind during completing writing tasks. What processes and how these processes take place inside learners' mind can be determined through completing a task and they are of utmost importance in defining, selecting, and sequencing the tasks which are appropriate for learners' levels in both second and FL learning settings. One of these processes which can play an important role in written language production is 'information processing'. From the information processing approach to task-based research, task complexity can be defined through intrinsic complexity (cognitive factors), perceived difficulty (learner factors), and task completion condition (interactional factors). The framework for defining cognitive task complexity adopted in this paper:

> distinguishes between dimensions of task complexity which can be manipulated to increase the conceptual and linguistic demands tasks make on communication, so creating the conditions for L2 'development', and the dimensions of task complexity which can be manipulated to increase the demands made on accessing a current interlanguage repertoire during real-time L2 'performance'. (Robinson, 2005, p. 5)

These two dimensions are discussed under 'resource-directing' and 'resource-dispersing' dimensions below.

1.1 Robinson's Cognition Hypothesis: Triadic Componential Framework

Robinson's Cognition Hypothesis (2005) and Skehan's Limited Attentional Capacity Model (Skehan & Foster, 1999, 2001) are two theoretical frameworks on which this study was based. Robinson's Cognition Hypothesis (2005), also known as Multiple Attentional Resources Model, states that human beings have unlimited attentional and memory resources which can be accessible whenever there is a need. The cognition hypothesis advocates the prediction that increasing cognitive task complexity which requires more attentional resources does improve language production qualities such as accuracy and complexity but not fluency. Robinson's Triadic Componential Framework embraces two dimensions dealing with cognitive loading, 'resource-directing dimensions', and 'resource-dispersing dimensions'. The former can be operationalized by whether the task requires learners to refer to events in the past or in the present, whether the task requires learners to refer to few or many elements, and whether the task requires learners to use spatial reasoning in completing writing task. On the other hand, the resource-dispersing dimensions deal with whether or not planning time or prior knowledge is given to learners and whether learners are required to complete one or multiple tasks simultaneously. In this study, the criterion used to label tasks (as 'topic', 'topic + idea', and 'topic

+ idea + discourse marker') is the amount of writing assistance which may require different cognitive loads on the writers, with less writing assistance requiring more cognitive load and changing the task into a more complex one.

1.2 Skehan and Foster's (2001) Limited Attentional Capacity Model

Another theoretical framework is Skehan and Foster's (2001) Limited Attentional Capacity Model. Unlike Robinson's model, Skehan and Foster's (2001) model proposes that all human beings have limited memory and attentional resources and when they are required to complete a cognitively demanding task, there will be some trade-off effects on different writing qualities (complexity, fluency, and accuracy). Its assumptions can be summarized as:

- Human beings have a limited information processing capacity; therefore, they prioritize some aspect(s) of language production over other ones.
- If a task demands a lot of attention to its content (more complex task), there will be less attention available to its language forms and vice versa.
- Learners prioritize the meaning and conveyance of it over its form during completing task if they are allowed to allocate attention freely (Van Patten, 1990).
-

1.3 Difference(s) between Robinson's Cognition Hypothesis and Skehan and Fosters' Limited Attentional Capacity Model

The first and foremost difference between Robinson's hypothesis and Skehan and Foster's model is that the former argues that learners can have access to multiple, unlimited, and non-competitional attentional and memory resources in completing a writing task, while Skehan and Foster reject it and focus on limited attentional resources. Cognition Hypothesis proposes that increasing cognitive demand of a task leads to less fluent, but more accurate and more complex language production because of humans' unlimited and non-competitional attentional resources. It means that if a task requires more attention to the content (meaning), it does not distract learners' attention from the form of language because there are enough memory resources available, but it has a negative effect on fluency of language production. The other area on which these two models diverge from each other is the prediction of the effect of increasing task complexity through resource-directing dimensions on language production quality. Whereas Skehan and Foster (2001) predict that increasing task complexity with respect to these factors leads to less fluent, less complex, and less accurate language production, Robinson (2005) argues that increasing task complexity with respect to these dimensions improves complexity and accuracy but reduces fluency.

Literature Review

Much of FL/L2 class time, particularly in school and university settings, is devoted to learning, teaching, and assessing writing skill (Benevento & Storch, 2011). Accordingly, research on writing tasks has attracted the attention of several scholars recently. There are several studies exploring

the effects of manipulating task complexity by the resource-directing factors on first and second language writing performance. Based on Robinson's (2005) Cognition Hypothesis and Skehan and Foster's (2001) Limited Attentional Capacity Model, Kuiken and Vedder (2008) conducted a study to explore the relationship between cognitive task complexity and linguistic performance in L2 writing. In their experiment, 91 Dutch university students of Italian and 76 students of French were required to complete two writing tasks with prompts of different cognitive complexity level. The measures of syntactic complexity, lexical variation, and syntactic accuracy provided support for the Cognition Hypothesis insofar as the written products of the more complex task came to be more accurate. Investigating a different resource-directing factor, Ishikawa (2006) examined the effects of manipulating task complexity with respect to the immediacy of time and place on 54 Japanese L2 learners' narrative writing. He reported that increasing task complexity with respect to the Here-and-Now dimension led to high level of accuracy, complexity, and fluency in learners' written language production.

Reviewing previous research makes it clear that there are only a few studies which examined the effects of resource-dispersing factors (e.g. planning time, number of tasks, and prior knowledge) on written language production. In their recent study, Ong and Zhang (2010) applied resource-dispersing dimensions of task complexity to detect the effects of task complexity on the fluency and lexical complexity of learners' argumentative writing. They manipulated task complexity using two factors: availability of planning time and provision of ideas and macrostructure. There were four groups to which different levels of planning time were given: extended pre-task, pre-task, and free-writing. One of these four groups was control group. Furthermore, the provision of the ideas and macrostructure had three levels: topic, ideas, and macrostructure group; topic and ideas group; and topic group. They found that increasing task complexity, with respect to the planning time continuum, resulted in significantly more fluency when it was measured by mean number of words produced per minute of the total time spent on the task and lexical complexity. Ong and Zhang (2010) also reported that the more complex task, through the provision of ideas and macrostructure, led to greater lexical complexity, but had no effect on fluency when measured by mean number of words produced per minute of transcription.

In a similar attempt, Ojima (2006) examined the effect of concept planning (as a resource-dispersing factor and as a form of pre-task planning) on three English as a Second Language (ESL) Japanese students' writing performance. He reported that pre-task planning produced greater fluency and complexity, but did not improve grammatical accuracy. In a similar vein, Wigglesworth and Storch (2009) conducted a study in order to determine whether there were any identifiable differences in the essays written by the learners working in pairs and those composed by the learners working individually. The essays were analyzed for fluency, complexity, and accuracy. Their findings revealed that collaboration had a positive effect on accuracy, but did not affect fluency and complexity of language production. In a recent study, Kormos (2011) investigated the effect of task complexity on linguistic and discourse features of narrative writing performance. He reported that FL participants produced more lexically complex texts. In addition, the findings indicated significant differences between L1 and FL narratives in terms of lexical variety, complexity, and syntactic complexity.

Whereas the effect of task complexity on oral language production has caught many researchers' attention in the past twenty years, there is considerably less research on how different complexity levels of task influence written output of FL learners. This study was therefore intended to first fill the gap in written task-based research. That is, this study attempted to explore the effects of increasing task complexity on English as a Foreign language (EFL) learners' argumentative writing using different amounts of writing assistance given to learners (topic, idea, and discourse marker). The second motivation for conducting this study was to see whether its results provide supportive evidence to Robinson' (2005) Cognition Hypothesis or to Skehan and Foster's (2001) Limited Attentional Capacity Model. Understanding whether there is any trade-off among

different writing qualities (e.g. lexical complexity, grammatical accuracy, and syntactic complexity) was still another aim of this study. More specifically, this research was meant to answer the following questions:

1. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect lexical complexity of EFL learners' argumentative writing?

2. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect fluency of EFL learners' argumentative writing?

3. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect grammatical accuracy of EFL learners' argumentative writing?

4. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect syntactic complexity of EFL learners' argumentative writing?

Method

Participants

Sixty upper-intermediate EFL learners (within the age range of 19-25) were recruited from two research sites, that is, three universities in Ardebil and Urmia, Iran, during the fall semester of 2011. They were selected from a pool of 90 learners. The writing section of an institutional Test of English as a Foreign Language (TOEFL) was used in order to homogenize the learners by excluding the outliers. The outliers were those who scored one standard deviation (SD) above and below the mean (M = 86 out of 100 and SD = 9). Thus, the scores which were below the 77 and above the 95 were considered as outliers and those which were within 77-95 took part in the main tasks. All of the participants had learned English in instructed academic setting. They had passed Advanced Writing course in university and at the time of the data collection some of them had just finished the third academic semester.

Materials

The writing section of an institutional TOEFL test was used to determine the general writing ability level of the participants. At this stage, the participants were required to write an argumentative essay evaluating advantages and disadvantages of human activities on the earth. Later on, three writing tasks with different amounts of writing assistance were given to the learners who were randomly assigned into one of these writing tasks. In the most complex writing task, the participants were invited to write an argumentative composition considering advantages and disadvantages of television on the relationships among family members and friends. In this group, only the topic of the writing task was given to the participants (topic group). In the medium-level complex task, the participants were invited to write the argumentative writing with the same topic as that of the first group. Some ideas were provided for this group, however (topic + idea group). The ideas encompassed two opposite points of view regarding the topic of argumentative essay. The topic of writing for the third writing task was the same as that of the previous two groups. In addition to topic and idea, some contrastive discourse markers were given to this group too (topic + idea + discourse marker group). This type of discourse markers is utilized dominantly in argumentative writing.

Procedure

First of all, in order to neutralize the possible effect(s) of language proficiency on the task completion procedure, EFL learners from three different universities in Ardebil and Urmia (Iran) took part in the study. The data was collected from students who had already passed an Advanced Writing course. Before the main writing task, participants were given the writing section of an institutional TOEFL in order to homogenize them in terms of their writing proficiency and to cross out the outliers. That is, before the experiment, 90 EFL learners completed a writing task in which they were asked to write an argumentative essay debating advantages and disadvantages of human activities on the earth. One of the researchers rated the writings based on the scoring rubrics offered by Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981). Following the researcher, a trained assistant rated approximately 20 percent of total essays, which were randomly selected. Inter-rater reliability, computed using Spearman rho, was very high between raters (ϱ = .96). After homogenizing the participants, 60 participants (out of 90) were randomly assigned to each of three main tasks in three different groups. The three groups were: (a) topic group; (b) topic + idea group; and (c) topic + idea + discourse marker group. Group 1 received only the topic of the writing (most complex task condition). Group 2 was provided with the topic and some relevant ideas for writing. Finally, group 3 received the highest amount of writing assistance, that is, topic, ideas, and some contrastive discourse markers (the least complex task condition). Based on the Triadic Componential Framework (Robinson, 2005), the researchers hypothesized that task complexity would increase incrementally from the topic, ideas, and discourse marker group, to topic and ideas group, to topic group. The participants were invited to write an argumentative essay debating advantages and disadvantages of watching television and its effects on relationships between friends and family members. They were required to write their composition in about 250-300 words within 90 minutes. A series of one-way ANOVAs (one independent variable [writing assistance] with 3 levels) were used to show if there were any significant differences among the three groups as far as lexical complexity, fluency, grammatical accuracy, and syntactic complexity of writings were concerned.

Participants' writings were coded in terms of lexical complexity, fluency, grammatical accuracy, and syntactic complexity. Different studies have used different measures to assess these different writing qualities. Ellis (2005, 2008) provides a fairly comprehensive list of such measures. He also points out that using multiple measures to assess each dimension of language performance may result in a more valid assessment. Thus, lexical complexity was measured through different procedures in this study, including the proportion of lexical words to function words (L/F), lexical density (LD), and mean segmental type-token ratio (MSTTR). The logic behind choosing the first two measures of lexical complexity was that, according to Halliday (1985) and Ure (1971), these measures are indices of the degree of orality versus literacy in both spoken and written discourse. They believe that the text which is more literate will be characterized by a higher degree of these measures of lexical complexity.

The criteria for classification of lexical and function words were based on Fontanini, Weissheimer, Bergsleithner, Perucci, and D'Ely (2005). In their definition, the function words are: modals, auxiliaries, determiners (articles, demonstratives, possessive adjectives, quantifiers, and numerals), pronouns, interrogative adverbs (what, when, how), negative adverbs (not, never), contracted forms of of pronouns, prepositions, conjunctions, discourse markers (but, so), sequencers (next, finally), particles (oh, uhm, well), lexicalized clauses (you know, I mean), quantifier phrases (anyway, somehow, whatever), lexical pause fillers (so, well), interjections (gosh, really, oh), and reactive tokens (OK, No!). Moreover, they defined the lexical words as nouns, adjectives, verbs, adverbs of time, place and manner, multiword verbs, idioms and contraction of pronouns, and main verbs (counted as one single item).

The second code of lexical complexity was lexical density (LD), which was calculated using the formula by Carter (1987):

$$\text{Lexical Density} = \frac{\text{number of separate (lexical) words}}{\text{total number of words in the text}} \times 100 \quad (\%)$$

The last measure of lexical complexity was MSTTR. Due to the fact that type-token ratio (TTR) is overly sensitive to sample size (Wolfe-Quintero, Inagaki, & Kim, 1998), one viable measure of lexical complexity which does not depend on text length (namely MSTTR) was run. According to Malvern and Richards (2002), MSTTR is an index that appears to have been originally recommended by Johnson (1994) and has been used in many other research studies. MSTTR truncates texts into sections of equal size and discards any remaining data. The TTR for each section is then recorded and the mean score of each section forms the final score. Section sizes are generally decided by the length of the smallest available text (Johnson, 1994). To find out the MSTTR in the present study, the students' written language productions were divided into segments of 117 words (the smallest available text in all three groups), the TTR of each segment was calculated and their average for the segments of written language produced by the students was calculated. Total number of different words (types) was divided by total number of words (tokens) in a text in order to calculate TTR in each truncated part. "For example, the phrase 'there is a woman who sits on a sofa' has a TTR of .88 because there are eight different words divided by nine total words" (Arslanyilmaz & Pederson, 2010, p. 387). According to Johnson (1994), a higher TTR is thought to indicate a greater lexical complexity. Thus, instead of using this raw type-token ratio, MSTTR was used in this study because TTR is a function of sample size, that is, larger samples of words will give a lower TTR because of less different words (Malvern & Richards, 2002).

Fluency was measured following the recommendations by Wigglesworth and Storch (2009). It was measured by: (a) total number of words (fluency I); (b) total number of T-units (fluency II); and (c) total number of clauses in each text (fluency III).

There are different measures for grammatical accuracy in task-based research. In this study, "to enhance both the validity of the assessments and the comparability of the results" (Ahmadian & Tavakoli, 2011, p. 48), some of the measures used by Wigglesworth and Storch (2009) were applied for measuring grammatical accuracy. In this study, grammatical accuracy was measured through the ratio of error-free Terminable units (T-units) to total T-units (EFT/T) and the ratio of error-free clauses to total clauses (EFC/C). A T-unit is an independent clause along with all subordinate and dependent clauses attached to or embedded in it, and it may be simple or complex sentence (Long, 1991; Kern, 1995). For instance, the sentence 'I ran down the stairs' consists of one T-unit, so is the sentence 'I ran down the stairs as fast as I could'. But a compound sentence is composed of more than one T-unit. For example, 'I ran down the stairs and the stairs twisted' has two T-units (Gaies, 1980; Ney, 1996). These measures of grammatical accuracy are textbook examples of global grammatical accuracy measures in task-based research. An error was operationalized in this study as any deviation in syntax, morphology, and lexical choice. Following Ellis and Yuan (2004), errors of punctuation, capitalization, lexical choice (e.g. kids vs. children) and spelling of any type were not taken into account unless they impeded the intended meaning.

Another writing quality measured in this study was syntactic complexity. Following Wolfe-Quintero et al. (1998), syntactic complexity was operationalized through two measures: the proportion of clauses to T-units (C/T) and the percentage of dependent clauses to total clauses (DC/C).

Data Analysis

To detect statistically significant differences among the three groups and to analyze the obtained data, a series of one-way between-groups ANOVAs (one independent variable [writing assistance] with 3 levels) was used. They were applied to detect whether manipulating cognitive task complexity had any effect on lexical complexity, fluency, grammatical accuracy, and syntactic complexity of EFL learners' argumentative writing.

Results

To answer the first research question: *Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect lexical complexity of EFL learners' argumentative writing?* a between-groups one-way ANOVA was used to discover the potential significant differences among the three groups for each measure of lexical complexity. Means and standard deviations of these measures are presented in Table 1.

Table 1

Means and Standard Deviations of Lexical Complexity in Three Groups

		N	Mean	Std. Deviation
L/F	TG	20	.90	.14
	TIG	20	.88	.14
	TID	20	.93	.13
	Total	60	.90	.14
LD	TG	20	46.92	8.46
	TIG	20	45.17	4.57
	TID	20	46.90	3.58
	Total	60	46.33	5.88
MSTTR	TG	20	66.06	4.00
	TIG	20	60.66	4.57
	TID	20	68.81	4.85
	Total	60	65.18	5.58

*TG: topic group; TIG: topic + idea group; TID: topic + idea + discourse marker group

The obtained results from one-way ANOVA showed that there were no significant differences among the three groups when lexical complexity was measured through L/F (F [2, 57] = .57, p = .56) and LD (F [2, 57] = .57, p = .56). In other words, manipulating task complexity (through providing different amounts of writing assistance) has no significant effect on lexical complexity measured by L/F and LD. Unlike these two measures, MSTTR, as another measure of lexical

complexity, found significant difference among the three groups (F [2, 57] = 17.06, p = .00). The effect size (.37) suggests that there is a large difference among the groups (Cohen, 1988) (see Table 2). The results of Post hoc Tukey test showed that the difference across the groups of participants does reach statistical significance for topic group and topic + idea group (p = .00), topic + idea group and topic + idea + discourse marker group (p = .00), but not for topic group and topic + idea + discourse marker group (p = .13). Hence, the first null hypothesis is rejected as far as MSTTR is concerned as a measure of lexical complexity. It is shown that the third group (with the highest amount of writing assistance) outperformed the others in this measure of lexical complexity.

Table 2

The Effect of Task Complexity on Lexical Complexity (L/F, LD, and MSTTR) (ANOVA)

		df	F	Sig.	Eta squared
L/F	Between Groups	2	.57	.56	.01
	Within Groups	57			
	Total	59			
LD	Between Groups	2	.57	.56	.01
	Within Groups	57			
	Total	59			
MSTTR	Between Groups	2	17.06	.00	.37
	Within Groups	57			
	Total	59			

Note: The mean difference is significant at the 0.05 level. * p <.05

To answer the second research question: *Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect fluency of EFL learners' argumentative writing?* one-way between-groups ANOVA was used to see whether there are any significant differences among the three groups in the measures of fluency. Descriptive statistics of the measures of fluency are showed in Table 3.

Table 3

Mean and Standard Deviations of Fluency in Three Groups

		N	M	SD
fluency I	TG	20	261.90	49.73
	TIG	20	233.85	66.51
	TID	20	289.25	66.25
	Average	60	261.66	64.46
fluency II	TG	20	20.35	4.19
	TIG	20	17.15	4.95
	TID	20	21.70	6.88
	Average	60	19.73	5.70
fluency III	TG	20	30.10	5.58
	TIG	20	27.60	8.41
	TID	20	35.45	9.25
	Average	60	31.05	8.44

The results of ANOVA indicated that there were significant differences among the three groups for all measures of fluency: fluency I (F [2, 57] = 4.07, p = .02, eta squared = .12), fluency II (F [2, 57] = 3.65, p = .03, eta squared = .11), and fluency III (F [2, 57] = 5.14, p = .00, eta squared = .15). The results suggested that increasing task complexity led to significant differences among the groups as far as all measures of fluency were concerned. Moreover, the results of Post hoc Tukey test showed that for all measures of fluency topic + idea group differed significantly from topic + idea + discourse marker group (fluency I [p = .01], fluency II [p = .02], and fluency III [p = .00]). Therefore, the second null hypothesis is rejected in this regard. Similar to lexical complexity, the third group outperformed the others in all measures of fluency.

Table 4
The Effect of Task Complexity on Fluency (I, II, III) (ANOVA)

		Sum of Squares	df	Mean Square	F	Sig.	Eta squared
fluency I	Between Groups	30693.23	2	15346.61	4.07	.02	.12
	Within Groups	214476.10	57	3762.73			
	Total	245169.33	59				
fluency II	Between Groups	218.43	2	109.21	3.65	.03	.11
	Within Groups	1701.30	57	29.84			
	Total	1919.73	59				
fluency III	Between Groups	643.30	2	321.65	5.14	.00	.15
	Within Groups	3563.55	57	62.51			
	Total	4206.85	59				

To answer the third research question: *Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect grammatical accuracy of EFL learners' argumentative writing?* a one-way between-groups ANOVA was used to provide a plausible answer to this question and to see whether there are any significant differences among the three groups in measures of grammatical accuracy.

Table 5
Means and Standard Deviations of Grammatical Accuracy in Three Groups

		N	Mean	Std. Deviation
Ratio of Error-free T-units to	TG	20	.76	.12
Total T-units	TIG	20	.59	.12
	TID	20	.77	.09
	Total	60	.71	.14
Ratio of Error-free clauses to	TG	20	.82	.11
Total Clauses	TIG	20	.65	.10
	TID	20	.84	.08
	Total	60	.77	.13

The results of one-way ANOVA for grammatical accuracy showed that there were significant differences among the three groups for both measures: EFT/T (F [2, 57] = 14.38, p = .00, eta squared = .33) and EFC/C (F [2, 57] = 19.30, p = .00, eta squared = .40). According to Cohen (1988), the effect sizes for both measures are large.

Table 6

The Effect of Task Complexity on Grammatical Accuracy (EFT/T and EFC/C) (ANOVA)

		df	F	Sig.	Eta squared
Ratio of Error-free T-units to Total T-units	Between Groups	2	14.38	.00	.33
	Within Groups	57			
	Total	59			
Ratio of Error-free clauses to Total Clauses	Between Groups	2	19.30	.00	.40
	Within Groups	57			
	Total	59			

The explored differences for EFT/T among three groups were significant between topic and topic + idea groups (p = .00) and topic + idea and topic + idea + discourse marker groups (p = .00). The difference is not statistically significant between the first and the third group (p = .94). Regarding EFC/C, the statistically significant differences among the three groups are allocated to topic and topic + idea groups (p = .00) and to topic + idea and topic + idea + discourse marker groups (p = .00). The difference between topic + idea and topic + idea + discourse marker is not statistically significant (p = .83). Regarding both measures of grammatical accuracy, the third group (the least complex writing task) outperformed the other two groups. Furthermore, it becomes clear that the second group was the weakest group in both measures of grammatical accuracy.

To answer the last research question: *Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect syntactic complexity of EFL learners' argumentative writing?* a one-way between-groups ANOVA was used to provide a plausible answer to this question and to see whether there are any significant differences among the three groups in each measure of syntactic complexity.

Table 7

Means and Standard Deviations of Syntactic Complexity in Three Groups

		N	Mean	Std. Deviation
C/T-units	TG	20	1.49	.18
	TIG	20	1.63	.31
	TID	20	1.68	.31
	Total	60	1.60	.28
DC/C%	TG	20	29.01	7.90
	TIG	20	46.25	10.95
	TID	20	39.42	12.80
	Total	60	38.76	14.57

The results of one-way ANOVA showed that there was no significant difference among the three groups when syntactic complexity was measured by the ratio of total number of clauses to total number of T-units (C/T), F (2,57) = 2.44, p = .09). They also made it clear that the observed difference among the three groups for the percentage of the ratio of dependent clauses to total clauses (DC/C) was statistically significant (F [2, 57] = 13.04, p = .00). The effect size is large (eta squared = .31) (Cohen, 1988).

Table 8

The Effect of Task Complexity on Syntactic Complexity (C/T and DC/C) (ANOVA)

		df	F	Sig.	Eta squared
C/T-units	Between Groups	2	2.44	.09	.31
	Within Groups	57			
	Total	59			
DC/C%	Between Groups	2	13.04	.00	
	Within Groups	57			
	Total	59			

The results of Post hoc Tukey test showed that the observed differences for DC/C among the three groups are significant between topic group and topic + idea group (p = .00) and topic group and topic + idea + discourse marker groups (p = .00). However, the difference is not statistically significant between the second and the third groups (p = .11). Regarding DC/C, the

second group outperformed the other two groups and the first (topic) group was the weakest group.

Discussion

This study was primarily aimed at examining the effects of task complexity on lexical complexity, fluency, grammatical accuracy, and syntactic complexity of EFL learners' argumentative writings. Regarding resource-dispersing factors, it was assumed that increasing task complexity resulted in language production which is less fluent, less complex, and less accurate. This prediction was based on Skehan and Foster's (2001) Limited Attentional Capacity Model and Robinson's (2005) Cognition Hypothesis. The obtained results for each of them will be discussed below.

As far as the first research question is concerned, it was found that the amount of writing assistance given to learners did not have any significant effect on lexical complexity when measured through the ratio of lexical to function words (L/F) and through lexical density (LD). There are some possible explanations for these findings. The first and foremost is the potential difference between lexical sophistication measures and lexical range measures. Ortega (1999) argued that previous research, which has reported significant results for lexical complexity measures, has focused mainly on lexical sophistication. These measures of lexical complexity (i.e. L/F and LD) consider lexical range measures in lieu of lexical sophistication measures and they are extremely sensitive to the length of text. Thus, observed results were not significant because all the texts were truncated into the same length (117 words). Another reason may be that, according to O'Loughlin (1995), these measures of lexical complexity (especially LD) are significantly affected by text types (description, narration, discussion, and role-play).

All of the three groups were required to write argumentative essays (the same text type); therefore, significant differences were not observed among them. Thus, these measures can be reliable indicators of text type and registeral appropriateness rather than lexical complexity of text (Halliday, 1985). Another explanation has to do with language proficiency level of participants. Because of their level of proficiency, the writers could not apply more sophisticated and lexical words than grammatical words. With respect to lexical complexity via MSTTR, the topic + idea + discourse marker group obtained significantly higher scores compared to topic and topic + idea groups. This finding can be explained by the fact that unlike L/F and LD, this measure of lexical complexity is not sensitive to the length of the text because it calculates the mean scores of type-token ratios in each truncated part. There are also some other explanations as to why the third condition allowed learners to write with greater lexical complexity (measured by MSTTR) compared to other two groups. One of them may be that participants in the least complex task might have been engaged in "deliberate and conscientious planning" (Ong & Zhang, 2010, p. 227) during completing task. That is, they were encouraged and had opportunity to write with an overall framework in mind. In other words, the writers in this group were somehow free from focusing on form and this allowed them to focus intensively on using more different and sophisticated words which led to higher scores of MSTTR. This lack of predetermined framework during writing task in the first and (to some extent) in the second groups might have decreased MSTTR scores in these groups. Another explanation for the enhanced lexical complexity may lie in the fact that the least complex tasks dispose extra cognitive information processing burden on the learners' mental capacities and lead them to use more different types of lexical items.

The findings of this study differed from those of Ong and Zhang (2010) who found that lexical complexity (measured via MSTTR) was higher in the task condition, where students were

provided with the least amount of assistance (most complex task) than in two other conditions. That is, they found that increasing task complexity through resource-dispersing factors (providing topic, idea, and macrostructure and planning time) resulted in more lexical complexity. However, the findings in the present study are in line with those of Ong and Zhang (2010), in measures of L/F and LD in that increasing task complexity produced no significant differences in lexical complexity measures.

In a similar vein, the lexical complexity findings of this study (via MSTTR) corroborate those of Kormos' (2011) study. In his study the task with given content (the least complex task) elicited higher lexical complexity as measured by the frequency of abstract words but no differences between tasks were found with regard to D-value.

The results of fluency measures in this study showed that giving the maximum amount of writing assistance did improve fluency of writing. In all fluency measures, the least complex task group outperformed the other two groups. This observation may be explained by one key reason. Writers in the third group (the least complex task) had enough room in their minds to dedicate for fluent language production. Moreover, discourse markers given to the third group led them to produce more compound structures and T-units (fluency I and II). Regarding fluency, this study was in line with that of Ahmadian and Tavakoli (2011), in that the opportunity to engage simultaneously in careful online planning and task repetition (the least complex tasks) enhanced fluency outstandingly.

On the other hand, the results of fluency measures diverge from those of Ong and Zhang (2010) who found that increasing task complexity by omitting the planning time led to greater fluency of writing. The results of the present study also differed from those of Ong and Zhang's (2010) study in which the group with the advantage of availability of the drafts during writing (the least complex group) did not produce more fluent language.

Regarding grammatical accuracy, findings have widely diverged in task-based research owing to different measures of grammatical accuracy. The results of this study showed that task complexity did lead to the production of texts which differ from each other from grammatical accuracy perspective. It was likely that, when working on the least complex task (with the more writing assistance), learners were able to put together their memory and cognition resources and paid their full attention to the form of their productions which, on the whole, led to more accurate language. The other reason for outperformance of the third group (topic + idea + discourse marker) might be that the learners in this group kept an eye on forms of language during writing when some key contrastive discourse markers were provided.

The results of measuring the grammatical accuracy (via the ratio of error-free T-units to total T-units (EFT/T) and the ratio of error-free clauses to total clauses (EFC/C)) are supportive results to both Robinson's Cognition Hypothesis (2005) and Skehan and Foster's (2001) Limited Attentional Capacity Model. They believe that increasing cognitive task complexity (along with the resource-dispersing factors) will lead to a decrease in the level of grammatical accuracy. This was the case in this study, that is, the least complex task group produced the texts which were more accurate than other two groups who were involved in more complex tasks.

Similarly, the findings of this study converge with those of Ahmadian and Tavakoli (2011). They reported that learners who enjoyed more careful on-line planning and task repetition (the least complex task) produced more accurate language than those who performed the task under pressured planning and without task repetition (the most complex task). Contrarily, the results of the current study contrast with those of the study by Kormos (2011) who manipulated the task complexity through the planning time. He found that the two tasks (± planning time) displayed

highly similar grammatical accuracy values. His results did not back up Skehan and Foster's (2001) Limited Attentional Capacity Model because the two tasks displayed highly similar accuracy scores. Mainly, the findings of the present study support those which have found a significant impact of task complexity on grammatical accuracy of L2 production when increasing task complexity results in less grammatically accurate language. This might be due to the assumption that providing ideas and contrastive discourse markers (which are of importance in argumentative essays) did channel attention toward specific features of the linguistic codes and structures; hence, it did speed up focusing on form.

Regarding syntactic complexity, the results showed that the least complex task group (topic + idea + discourse marker group) outperformed other two groups when it was measured by the ratio of dependent clauses to total clauses. This observation can be explained by the fact that, according to Skehan and Foster's (2001) model, increasing task complexity draws away learners' attention from focusing on form of language to conveying the message of content. Thus, the writers in the third group which were provided by discourse markers produced more complex sentences with the dependent clauses embedded in them. Conversely, the measures of syntactic complexity via the ratio of total clauses to total T-units did not reach the significant level. The results of this study converge with those of Storch (2005), in that his study found that learners who wrote in pairs produced better texts in terms of task fulfillment, grammatical accuracy, and complexity.

The present study, in sum, showed that manipulation of task complexity affects fluency, grammatical accuracy, and syntactic complexity (via DC/C), but not lexical complexity (except when measured via MSTTR). Thus, the findings did not provide strong evidence in support of all the predictions made by Skehan and Foster's (2001) Limited Attentional Capacity Model and Robinson's (2005) Cognition Hypothesis as far as lexical and syntactic complexity were concerned.

Conclusion and Implications

This study attempted to investigate the effects of cognitive task complexity on different writing qualities: lexical complexity, fluency, grammatical accuracy, and syntactic complexity in EFL learners' argumentative writing. Having identified a gap in task-based research in written discourse, this study attempted to bridge this gap by investigating the effect of cognitive task complexity on learners' writing performance. This was done on four different but somehow related writing qualities in three task conditions (with different amounts of writing assistance). The results of the measures of lexical and syntactic complexity provided supportive evidence (although partially) for both Robinson's and Skehan and Foster's models. In the same vein, the results of fluency and grammatical accuracy measures contributed supportive evidence to both above-mentioned models.

The results of this study also provided support to Skehan's (1998) trade-off Hypothesis when the trading was between task complexity and different writing qualities. But, there were no such trading-off effects among lexical complexity, fluency, grammatical accuracy, and syntactic complexity.

The present study has a number of theoretical and pedagogical implications for SLA researchers, teachers, syllabus and task designers, and language testing specialists. The first and the most important is that the nature of the different processes and specially information process (involved during completing task) can be clearly established in SLA settings. As Ellis (2009) argued, task-

based language teaching (TBLT) is of utmost importance for language learning and teaching since it operationalizes the theory of SLA, makes it more tangible, and provides more insightful perspective for learning and teaching a language (especially foreign language). A valid and crucial criterion for designing, selecting, grading, and sequencing pedagogical tasks is in forefront of teachers' and task designers' attention. Therefore, the findings of this study can be regarded as practical basis for above-mentioned purposes. Moreover, the findings of the current study suggested that focusing on cognitive capabilities of the learners as well as the cognitive load of the structure of the task is of more importance in language teaching and learning than in any other settings.

In spite of some useful findings which can be considered as explicit basis for writing assessment, some limitations need to be acknowledged. One of the limitations may be that this study did not reveal other important effects of decreasing task complexity on written outputs because only 20 participants were recruited for each writing tasks. The other limitation which may be in the centre of future research attention is that, ask complexity in the present study was manipulated through only resource-dispersing dimensions (using Robinsons' Triadic Componential Framework terminology), and the other dimensions, that is resource-directing factors, were not taken into account. To fully detect such effects on writing performance, it would be insightful to manipulate task complexity by both types of dimensions.

To overthrow the mentioned limitations and to obtain enlightening results, some recommendations for future researchers are made to ponder on. First and foremost, future research may raise the number of the participants in each writing task in order to be a good indicator of task complexity effects on the participants' writing performance. The other implication for further research may be that they may "need to tackle different task complexity challenges which may have an impact on the quality of EFL writing" (Sadeghi & Mosalli, 2012, p. 63). It means that other types of tasks (story narrating or story making) in other genres (descriptive or expository) by different definitions of task complexity should also be investigated. Last but not least, further research may focus exhaustively on other writing qualities such as authenticity and appropriateness of discourse.

References

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35-59.

Arslanyilmaz, A., & Pedersen, S. (2010). Improving language production using subtitled similar task videos. *Language Teaching Research, 14*(4), 377-395.

Bae, J., & Bachman, L. F. (2010). An investigation of four writing traits and two tasks across two languages. *Language Testing, 27*(2), 213-234.

Benevento, C., & Storch, N. (2011). Investigating writing development in secondary school learners of French. *Assessing Writing, 16,* 97-110.

Carter, R. (1987). *Vocabulary: Applied linguistic perspectives*. London: Rutledge.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in second language* (pp. 3-34). Amsterdam: John Benjamins.

Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production. *Applied Linguistics, 30*, 474-509.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*, 59-84.

Fontanini, I., Weissheimer, J., Bergsleithner, J. M., Perucci, M., & D'Ely, R. (2005). Memória de trabalho e desempenho em tarefas de L2. *Revista Brasileira de Lingüística Aplicada, 5* (2), 189-230.

Gaies, S. J. (1980). T-unit analysis in second language research: Applications, problems and limitations. *TESOL Quarterly, 14*(2), 85-98.

Halliday, M. A. K. (1985). *Spoken and written language*. Oxford: Oxford University Press.

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.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20,* 148-161.

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.

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL compositions: A Practical approach.* Rowley, MA: Newbury House.

Johnson, W. (1994). Studies in language behavior: I. A program of research. *Psychological Monographs, 56,* 1-15.

Kern, R.G. (1995). Restructuring classroom interaction with network computers: Effects on quantity and characteristics of language production. *The Modern Language Journal*, *79*, 457-76.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective*, (pp. 39-52). Amsterdam: John Benjamins.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity, *Language Testing, 19* (1), 85-104.

Ney, J. W. (1996). Review of the grammatical structures written at three grade levels. *Language Learning, 16* (2), 230-235.

Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers. *System, 34*, 566-585.

O'Loughlin, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language Testing, 12,* 217-237.

Ong, J., & Zhang, L., J. (2010). Effects of task complexity on fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, *19*, 218-233.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21*, 109-148.

Robinson, P. (2005). Cognitive complexity and task sequencing: A review of studies in a Componential Framework for second language task design. *International Review of Applied Linguistics, 43*, 1-33.

Robinson, P., Chi-chien Ting, S., & Urwin, J. (1995). Investigating second language task complexity. *RELC, 26,* 62-78.

Sadeghi, K., & Mosalli, Z. (2012). The effect of task complexity on fluency and lexical complexity of EFL learners' argumentative writing. *International Journal of Applied Linguistics & English Literature 1*(4), 53-66.

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

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49,* 93-120.

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.

Storch, N. (2005). Collaborative writing: product, process, and students' reflections. *Journal of Second Language Writing, 14,* 153-173.

Ure, J. (1971). Lexical density and register differentiation. In G. E. Perren, & J. L. M. Trim (Eds.), *Applications of linguistics: Selected papers of the second international congress of applied linguistics* (pp. 443-452). Cambridge: Cambridge University Press.

VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, *12*, 287-301.

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing, 26*(3), 445-466.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.Y. (1998). *Second language development in writing: Measures of fluency accuracy and complexity*. Honolulu, HI: University of Hawaii Press.

Karim Sadeghi has a PhD in TEFL/TESOL (Language Testing) from the University of East Anglia in Norwich, UK. His main research interests include: language testing, alternative assessment, and critical research evaluation. He is currently the editor-in-chief and managing editor of *Iranian Journal of Language Teaching Research*. In addition to serving as an editorial and advisory member of national and international journals, he is also an advisory board member of *International Encyclopedia of ELT* (edited by John Liontas, University of South Florida).

Zahra Mosalli is an MA holder of ELT from University of Urmia. She has published some research papers in different journals (IJALEL, WASJ). She has also participated in two international conferences: ICALT and TELLSI 10th. She has taught English language in English language institutes and in Mohagheghe Ardebili University. She is interested in research on SLA, task-based language teaching and learning, and language assessment.