

The Effect of Task Complexity on the Quality Narrative Writing: Case of Male and Female Iranian EFL learners

By

¹*Afsaneh Ebadi* and ²*Davoud Amini*

Islamic Azad University, Ahar Branch, Islamic Azad University, Ahar, IRAN.

Assistant professor, Azarbaijan Shahid Madn University, AZARBAIJAN.

Abstract

This study attempted to explore the effect of manipulating task complexity on narrative writing quality of male and female Iranian EFL learners in terms of lexical complexity, fluency, grammatical accuracy, and syntactic complexity. Task complexity was manipulated through applying resource-dispersing dimensions. All 60 participants (30 males and 30 females) 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. One-way ANOVAs were 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. 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.

Keywords: *task complexity; lexical complexity; fluency; grammatical accuracy; syntactic complexity; narrative writing, Gender*

1. Introduction

As Ozyildirim(2010) proposed learners in EFL classrooms can not become successful writers and they receive poor grades on their writings, not because they lack the ability to communicate , but because they can not seem to manage their time and texts properly.

Daniel kies (1995) notes that narrative writing is a difficult skills for EFL learners in EFL classrooms ,and also notes that EFL learners don't know where to begin, how to organize the materials and put things in unsuitable places, so, it makes their products meaningless, it can be said that there is a need to find some new activities in order to increase the learners performances in this crucial skill.

Perhaps the most difficult aspect of writing among Iranian foreign language learners is that it is always accomplished and assessed through a series of blurred and peddling methods. So there should be some valuable methods in order to improve and assess the quality of the learners' writing performance. One of the methods is considering the effect of task complexity on second language learners' writing performance.

There are different but similar definitions of task complexity. As Gilabert (2004) asserts, the need for sequencing tasks from simple to complex in a reasoned way that will foster interlanguage development was the impetus to the emergence of the concept of task complexity. According to Robinson (2005), task complexity is defined as "the result of intentional, memory, reasoning, and other information processing demands imposed by the structure of the task on language learner." Skehan (1998) uses the term interconnectedness to refer to complexity: more elements or characters make for greater task difficulty.

On the other hand, although there are large numbers of publications related to task-based language learning, teaching, and testing, (e.g. Ellis, 2003; Gilabert, 2004; Kuiken & Vedder, 2008; Skehan, 1998;),

there have been few studies to investigate the effects of task complexity on L2 learners' narrative writing performance. Thus, the present study set out to investigate the effects of task complexity on L2 learners' narrative writing performance in terms of accuracy, fluency, and complexity. It is hypothesized that task-complexity has some positive or negative impact on learners' written performance in terms of accuracy, fluency, and complexity. Among a variety of techniques that have been used on improving students' writing skills, just a few of them have been devoted to investigate the effect of task complexity in the realm of resource-dispersing dimension on EFL learners' writing. So, this study looks at this dimension of task complexity.

Moreover, this study is genre-specific, that is, it tries to focus on narrative genre of writing which is new in itself. So, this research could be used as a sort of highly effective study to develop EFL learners' narrative writing quality.

Research Questions

1. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect lexical complexity of EFL learners' narrative writing?
2. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect fluency of EFL learners' narrative writing?
3. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect grammatical accuracy of EFL learners' narrative writing?
4. Does task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) affect syntactic complexity of EFL learners' narrative writing?
5. If the answers to three first research questions is yes, does task complexity affect syntactic, lexical, fluency and grammatical complexity of male and female learners differently?

Research Hypotheses

H₀₁: Task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) has no effect on lexical complexity of EFL learners' narrative writing.

H₀₂: Task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) has no effect on fluency of EFL learners' narrative writing.

H₀₃: Task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) has no effect on grammatical accuracy of EFL learners' narrative writing.

H₀₄: Task complexity (topic only vs. topic + idea vs. topic + idea + discourse marker) has no effect on syntactic complexity of EFL learners' narrative writing.

H₀₅: Task complexity does not affect syntactic, lexical, fluency and grammatical complexity of male and female learners differently?

2. Methodology

Research Design

According to Mackey and Gass (2005) the design of this research was a comparison group design, participants were randomly assigned to three groups. It is a subcategory of quasi-experimental design.

Participants

Sixty advance EFL learners (30 males and 30 females) within the age range of 19-45 were recruited from two research sites in Ardebil, Iran, during the fall semester of 2014. 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 more than one standard deviation (SD) above and below the mean ($M = 86$ out of 100 and $SD = 9$) and they

were assigned randomly (simple randomization) into three groups of topic (n = 20), topic+ idea (n= 20), and topic + idea + discourse markers (n = 20).

Instruments

All instruments which were used in the study are going to be described in more details in the following section.

Pre-test

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 essay evaluating advantages and disadvantages of human activities on the earth.

Post-test

Both pre- and post-test were essay writing. Learners were required to write their composition in about 250-300 words within 90 minutes. They were adapted by the teacher and prior to the actual research they were piloted in order to estimate the allotted time to complete the task or to improve any possible error or misunderstanding.

Procedure

First of all, in order to neutralize the possible effect (s) of language proficiency on the task completion procedure, EFL learners from Sarab University in Ardebil, 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, 60 EFL learners completed a writing task (pre-test) in which they were asked to write an essay debating advantages and disadvantages of human activities on the earth and the researcher 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, was very high between raters ($\rho = .96$). After homogenizing the participants (based on the pre-test) 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 (TG); (b) topic + idea group (TIG); and (c) topic + idea + discourse marker group (TID). 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 a narrative 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. 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 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): $\text{Lexical Density} = 100 (\%) \text{ total number of words in the text divided by number of separate (lexical) words.}$

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 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).

3. Data Analyses and Results

In order to ensure that the learners in three groups enjoyed the same level of language proficiency prior to the main study, their scores on the pretest were put into a One-way ANOVA (Table 3.1) and the results revealed that there were no significant differences across the participants on their level of proficiency ($p = 0.80 > 0.05$).

Table 3.1: One-Way ANOVA for Scores on Pre-test

Source	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	0.358	2	0.179	0.212	0.809
Within Groups	48.075	57	0.843		
Total	48.433	59			

The obtained results from one-way ANOVA (Table 3.2) 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 ($F [2, 57] = .57, p = .56$). 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 3.2: 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			

The results of ANOVA (Table 3.3) indicated that there were significant differences among the three groups for all measures of fluency: fluency I ($F [2, 57] = 4.07, p = .02, \eta^2 = .12$), fluency II ($F [2, 57] = 3.65, p = .03, \eta^2 = .11$), and fluency III ($F [2, 57] = 5.14, p = .00, \eta^2 = .15$). The effect size (.15) suggests that there is a large difference among the groups (Cohen, 1988). 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 Turkey 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 3.3: 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				

The results of one-way ANOVA for grammatical accuracy (Table 3.4) showed that there were significant differences among the three groups for both measures: EFT/T ($F [2, 57] = 14.38, p = .00, \eta^2 = .33$) and EFC/C ($F [2, 57] = 19.30, p = .00, \eta^2 = .40$). According to Cohen (1988), the effect sizes for both measures are large.

Table 3.4: 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	
within Groups	57			
Total	59			

The results of one-way ANOVA (Table 3.5) 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 ($\eta^2 = .31$) (Cohen, 1988).

The results of Post hoc Turkey 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.

Table 3.5: The Effect of Task Complexity on Syntactic Complexity (C/T and DC/C) (ANOVA)

	Df	F	Sig.	Eta squared
C/T				
Between Groups	2	2.44	.09	.31
Within Groups	57			
Total	59			
D/C%				
Between Groups	2	13.4	.00	.35
Within Groups	57			
Total	59			

As there are two independent variables, namely task complexity and gender, the present data analysis utilized a two-way ANOVA to explore the impact of sex and type of task complexity on learners' writing skill. As can be seen in Table 3.6 the p-value of gender was higher than 0.05 ($.50 > .05$). It indicates the fact that there was no significant main effect for gender and gender had not a significant effect on students' writing accuracy..

Table 3.6: Two-way ANOVA for Scores on Post-test

Source	SS	Df	Mean Square	F	Sig.	Partial Eta Squared
Group	26.585	2	13.293	34.656	.000***	.562
Gender	.171	1	.171	.445	.508	.008
Group*Gender	.329	2	.165	.429	.653	.016
Error	20.712	54	.384			
Total	16369.600	60				

4. Discussion of Findings

Teachers can use new techniques to enhance their learners' performances in writing narratives which was considered throughout this work. Therefore, finding out the useful class activities that can be used in order to increase learners' abilities to write more reasonable compositions, can help the teachers to reach acceptable outcomes in their teaching classrooms and causes the learners improvement in writing skill particularly in EFL settings. The present study attempted to evaluate the effect of task complexity on the quality of EFL learners' narrative writing.

This study was aimed at examining the effects of task complexity on lexical complexity, fluency, grammatical accuracy, and syntactic complexity of EFL learners' narrative 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 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.

References

- Darayesh,A.(2003).The Effect of a *Proposed Program Based on semantic mapping and Brainstorming Strategies on developing the English Writing Ability and Attitudes of the First Scientific Secondary Students*, Unpublished Ph.D. Thesis. Amman Arab University for Gratitude Studies, Amman ,Jordan
- Ellis,R.(2003).*The Study of Second Language Aquisition* .oxford:Oxford University Press
- Gilabert, R. (2004). *Task complexity and L2 narrative oral production*. *Unpublished doctoral dissertation*, Universitat deBarcelona
- Ozyildirim, I (2009). Narrative analysis: An analysis of oral and written strategies impersonal experience narratives. *Journal of pragmatics*, 41, 1209, 1222, Retrieved on 2012,3,14 from [Http://dy.doi.org/10.1016/J/pragma.2009.01.003](http://dy.doi.org/10.1016/J/pragma.2009.01.003)
- Ozyildirim, Ataman. (2010). A more Timely and Useful Index of Learning Indications. Instructional working papers.03-01, The Conference Board, educational program
- 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. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45(1), 99-140.
- Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-57.
- Robinson, P. (2003). The cognition hypothesis, task design and adult task-based language learning. *Second Language Studies*, 21(2), 45- 105.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43, 1–32. 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). Cambridge: Cambridge University Press.
- Yu. H. (2005) An Exploratory study of narrative structures in china tertiary EFLlearners monologic production. Beiying, *Foreign language teaching and research press*