

The Effects of Concept Mapping Strategy and Aural vs. Written Prompts on Writing Test Performance under Different Planning Conditions

N. Farshi *
M.A., TEFL
University of Isfahan
email: najmeh.farshi@gmail.com

M. Tavakoli
Associate Professor, TEFL
University of Isfahan
email: mr.tavakoli14@gmail.com

Abstract

This study examined the effects of aural and written prompts under two planning conditions (i.e. pre-task planning and no planning) on complexity, accuracy, and fluency of test takers' writing production. Forty learners in an English institute, who had already been classified as intermediate according to the Oxford Placement Test, were assigned to two planning conditions (i.e. no planning and pre-task planning). Then the planning groups were further divided into another two groups: with aural prompt and with written prompt. Also, concept mapping strategy was applied during pre-task planning time by the test takers. The results obtained from t-test and two-way ANOVA revealed that the candidates who received the written prompt utilized their planning time better and produced more fluent written texts than those who received the aural prompt. Furthermore, neither concept mapping strategy with aural prompt nor concept mapping strategy with written prompt led to more complex or more accurate writings. Finally, the interaction of no planning condition and written prompt had a significant effect on complexity in comparison with the pre-task planning condition with written prompt. Also, written prompt under no planning condition had a significant effect on complexity in comparison with the same planning condition with aural prompt. It was concluded that the planned conditions, concept mapping strategy and the received prompts had little effect on the test takers' writing performance.

Keywords: aural prompt, concept mapping, planning, testing context, written prompt

Received: 10/19/2013 Accepted: 05/28/2014

* Corresponding author

1. Introduction

In testing and pedagogic contexts planning is manifested differently. In a testing context, test takers' awareness of their being assessed may direct their attention to accuracy and divert their attention away from fluency and complexity. Attention to accuracy during careful online planning may blunt the effect of any strategic planning on complexity, accuracy, and fluency (Ellis, 2005). In this context planning generally has led to no favorable results (e.g., Elder & Iwashita, 2005; Elder, Iwashita & McNamara, 2002; Iwashita, McNamara & Elder, 2001; Wigglesworth, 2000; Wigglesworth & Elder, 2010). In pedagogic contexts, on the other hand, several studies on the effects of strategic planning on language production have shown fluency enhancement with more mixed results for complexity and accuracy (Ellis, 2009; Foster & Skehan, 1996; Ortega, 1999, 2005; Sangarun, 2005; Skehan, 1998; Skehan & Foster, 1997; Wendel, 1997; Wigglesworth, 1997; Yuan & Ellis, 2003). In testing context, test takers try to avoid errors rather than to use all of their language resources (Wigglesworth & Elder, 2010). Another probable reason for the ineffectiveness of pre-task planning in testing context is less planning time provided by testing studies than the planning time provided by classroom ones (Ellis, 2009). In testing situation, for practical reasons, it is not possible to provide such ample planning time as given in pedagogic context.

In this regard, understanding the trade-off among complexity, accuracy, and fluency, as three aspects of language production, seems necessary. Skehan (2009) states that attending to all three aspects, at the same time, is not easy. This is because human's attentional resources are limited and therefore devoting attention to one area can decrease attention to others. Thus, planning can make tasks easier by removing the time pressure and buying time for processing (Skehan, 2009). Robinson (2001) on the other hand, has the opposite idea. He sees a correlation between complexity and accuracy. Planning according to his research makes the tasks less demanding and therefore less effective. When the subjects find the task which is made easier by planning, they may ignore forms and seek to invest on meaningful production. Studies on planning in general have elaborated on the issue of trade-off between form and meaning (e.g., Ahmadian & Tavakoli, 2010; Ellis & Yuan, 2004; Foster & Skehan, 1996; Foster & Skehan, 1999; Yuan & Ellis, 2003).

Up to now, most of the planning studies have been concerned with investigating the effect of planning on oral production. Very few studies

have developed this issue for writing performance. For example, Ellis and Yuan (2004) designed their research based on the planning studies of oral performance and Kellogg's (1996) writing model.

Although writing processes bear strong likeness to speaking processes (Ellis & Yuan, 2004), they need longer time for both pre-task and online planning; therefore, the students can play havoc with the extra time provided for planning, or they may take this opportunity to remove their mistakes. Therefore, concept mapping strategy can be effectively instructed for advance planning (Ojima, 2006). Strategy can help the students to organize their ideas in a 'network of relationships' (Chularut & Debacker, 2003, p.249) and link the received information to their previous knowledge (Ausubel, 1963).

Therefore, the purpose of this study was to determine how pre-task planning, with the strategy of concept mapping applied during planning time in testing context, would affect EFL learners' writing test performance in terms of complexity, accuracy and fluency.

2. Theoretical Background

2.1 Planning in testing context

What makes the most obvious difference between a testing context and a classroom context is how performing a task is judged. In a testing context, it is of utmost importance to perform a task properly in comparison with a classroom or laboratory context (Ellis, 2005).

In addition to the participants' language competence, a whole range of factors can affect language task performance in a testing context. How much a testee employs performance strategies and the testing conditions under which the task is done are among these factors (Tavakoli & Skehan, 2005).

In contrast to the generally positive results of planning in pedagogic context, studies of planning in testing context have revealed a lot of inconsistencies. For example, Wigglesworth (1997) who examined the effect of pre-task planning for simple and more difficult tasks on oral test performance of high and low proficiency candidates, found no significant difference in the raw scores as a function of planning time or lack of planning time. Besides, high-proficiency candidates benefited from the presence of planning time in terms of complexity but this was not the case for the low-proficiency candidates. For both the high-proficiency and the low-proficiency candidates, on the easier tasks, planning time did not make any difference. In another study by Tavakoli and Skehan (2005), strategic

planning, task structure, and proficiency were the independent variables manipulated to see their effects on learners' narrative test performance. With regard to planning, the subjects were divided in half, performing their tasks under pre-task planning and online planning condition. The results clearly demonstrated that pre-task planning brought about significant changes in performances in terms of the triad of complexity, accuracy, and fluency.

Wigglesworth (2000) focused on the conditions that can damagingly make oral assessment tasks more difficult for language learners. She systematically manipulated structure and familiarity as the task characteristics, and identified nativity or non-nativity of the interlocutors and variation of planning time as task conditions. In fact, planning time was manipulated with structure and with task familiarity interactively. As a result, structured tasks were easier than unstructured ones, but planning time made both of them more difficult. Also, familiar tasks were easier than unfamiliar ones but planning time increased familiar tasks' difficulty and was neutral to unfamiliar activities. The reason for the adverse influence of planning time according to the study can be its promotion of more complicated ideas without the test takers being able to change these ideas into accurate and fluent linguistic output.

Iwashita, McNamara, and Elder (2001) varied narrative tasks perspective, immediacy, adequacy, and planning time, in order to change task demands and consequently the complexity, accuracy, and fluency of the candidates' performance. They predicted that lack of planning time would increase task difficulty. According to the results, no effect of performance conditions, task versions, or the interaction of performance condition and task versions on fluency and complexity was found. Only for the immediacy dimension, when the candidates narrated the story with the pictures in front of them, a significant influence on accuracy was derived. One of the possible reasons for the insignificant results was the big difference between pedagogic and testing context, since the latter can alter the participants' cognitive focus on task accuracy regardless of the conditions or demands. Moreover, the importance of the results in testing context could cause the candidates to be too inhibited to be able to produce complex language. In other words, they preferred to confine themselves to the boundaries of their believed knowledge. In line with this investigation, Elder and Iwashita (2005) tried to gain more knowledge about the provision of strategic planning time in testing context. Their participants were asked to complete eight narrative tasks in which planning time was manipulated. Based on

previous research especially the one by Skehan (1998), it was hypothesized that giving planning time to the test takers would make the tasks easier and would lead to more complex, accurate, and fluent responses. Although the results revealed that no significant differences were found between the two planning conditions (pre-task planning and no planning), the participants had recognized the planned tasks easier but at the same time less enjoyable than the no planned ones. Elder and Iwashita (2005) proposed several probable reasons for the outcomes, among them were the unfamiliarity of the test takers with how they could benefit from the available planning time, simple narratives which may not require complex wording, and generous online planning time for task completion. But specifically they mentioned the inherent difference of assessment with classroom context in what they called as language behavior.

On the other hand, Wigglesworth and Elder (2010) studied the effect of the interaction of planning, proficiency, and task, as independent variables, on IELTS oral section performance. They intended to check the variety of performances, in terms of raw scores and discourse quality, under variation of pre-task planning time. Investigation of students' perceptions and attitudes toward planning and their planning strategies was also importantly included in the purposes of the study. Therefore, 90 candidates who were divided into advanced and intermediate levels, performed three parallel tasks under 0 min, 1 min, and 2 min strategic planning time. With regard to the candidates' raw scores, and the discourse measures of complexity, accuracy, and fluency, the effects of task or planning time were not significant. With regard to the interpretations and use, most of candidates highly admired being given the planning opportunity for its organizing facilitation, but the fact was that the testing context made them too anxious to be able to use their planning chance. Also, utilization of the strategies by the candidates was significantly better than when planning time was not permitted although the number, the type of the strategies, or the amount of planning time made no significant differences in their performance. According to the researchers, memory limitations that lead to plan only the first few utterances, the 0 min planning condition in which online planning was possible, or the different inferences of the candidates and the raters of the desired speaking task, could also be responsible for the insignificant results. Despite the null findings, Wigglesworth and Elder (2010) state that inclusion of planning time does not seem to be superfluous because of the candidates' preference expressions and also face validity reasons.

Ellis (2009) points to a probable reason for ineffectiveness of assessment task planning that is shortage of task completion time in this setting compared with classroom or laboratory ones. With regard to Ellis's point of view and since writing takes longer online time than speaking does, writing tests were conducted rather than oral ones in order to see whether planning would be beneficial in writing contexts. Ellis and Yuan (2004) who investigated the effect of pre-task and online planning on complexity, accuracy, and fluency of learners' second language narrative writing in classroom context, considered Levelt's (1989) speaking model and Kellogg's (1996) writing model as counterparts. They found that pre-task planners engaged in formulation (which is parallel to speech conceptualization) of their writings more than online planners and no planners. Consequently, they produced more fluent language than the other groups. Online planners, in contrast, produced more accurate writings which can be assigned to the candidates' online monitoring of their outcome.

To sum up, review of the planning studies in testing context shows that none of them have been done with reference to writing performance. This article was aimed to look at pre-task planning in testing context from writing perspective. Following Ellis and Yuan's (2004) creativity in relating Kellogg's (1996) model of writing with different types of planning, this article applied this model to writing production of pre-task planners in testing context. But, based on another proposal by Ellis (2005, p.26) in which a testing context is seen through a 'psychological context' regardless of modality, it was hypothesized that strategic planning of the writing tests would be of no significant benefit.

2.2 Concept mapping strategy

Concept mapping was originally developed in Cornell University as a graphical representation of knowledge which is elaborated on while being engaged in creative thinking (Novak, 1992). Relevant to the principles of schema theory, it combines foreknowledge with new information (Ojima, 2006); therefore, it facilitates analyzing and synthesizing information (Novak, 2010). This can assuage testers' anxiety and make their performance more satisfactory (Liu, 2011).

Trying to plan to write, writers can form an internal and abstract representation of their knowledge by depicting it as appropriate keyword diagrams. Combined into thoughts, these concept illustrations ease the process of writing (Pieronek, 1994).

Schultz's (1991) study on the influence of concept mapping strategy on L2 writing showed the students' improved writing abilities and their being motivated to participate in the discussion and suggested the strategy of concept mapping as a useful pre-writing performance tool. Liu (2011) also checked the effect of concept mapping strategy in pre-writing phase on learners' writing performance. Ninety-four participants received no-mapping, individual-mapping, and cooperative mapping as different types of treatment. Being divided into three language levels (i.e. high, middle and low), the learners were assigned three writing tasks and had to handle them within nine weeks. The results revealed that individual and cooperative mapping proved useful for low and middle-level learners. Furthermore, high-level learners with the individual-mapping treatment performed their tasks significantly better than their counterparts with two other kinds of treatments. Generally, the concept maps had provided the chance of visualizing the ideas and noticing what was irrelevant to the main topic.

The influence of concept mapping as a form of pre-task planning on learners' writing performance was also investigated by Ojima (2006). In this study, three Japanese students were to write with and without the application of concept maps and were asked about their writing experience through questionnaires and interviews. Complexity, accuracy, and fluency of their texts revealed the positive effect of pre-task planning on the learners' performance. Moreover, learners' feedback on their writing tasks reflected their unique challenge using the strategy.

Research on concept mapping as a language planning strategy is very limited, and most of the studies have been done in L1 writing contexts. To date no study has been done on the effect of concept mapping as a pre-task planning strategy in writing from a testing perspective. This study was aimed at filling this gap.

3. The Present Study

The study, which was conducted in testing context, was a between-group quasi-experimental design aimed to investigate the effects of aural and written prompts on complexity, accuracy, and fluency of EFL learners' writing test performance under different planning conditions.

Building on the above-mentioned theoretical and empirical rationales, the following research questions were formulated:

RQ1: Does concept mapping strategy through aural and written prompt significantly affect test takers' writing production in terms of complexity, accuracy, and fluency?

RQ2: Do no planning and pre-task planning with aural and written prompts significantly affect test takers' writing production in terms of complexity, accuracy, and fluency?

Null Hypothesis: In light of the literature reviewed in the previous section, it is hypothesized that planning conditions and concept mapping strategy with aural and written prompts have no significant effects on test takers' writing production in terms of complexity, accuracy and fluency.

4. Method

4.1 Participants

Forty intermediate level test candidates both male and female, who had joined an English course in a language institute in Iran, participated in this study voluntarily. The study was conducted in four classrooms, and each classroom had about ten students. According to the regulations of the language institute and its placement test (the Oxford Placement Test with essential characteristics of validity and reliability), the students of these classrooms had already been classified as intermediate. This research was conducted under a natural classroom situation. Therefore, the treatment was randomly assigned to the intact classes. Actually, the participants were assigned into two planning condition groups with each group being further divided into another two groups. The schematic representation of the participants of the four groups (regarding aural and written prompts and planned conditions) is depicted in Table1:

Table 1. Number of participants

| Planning condition | No Planning (NOP) | | Pre-Task Planning (PTP) | |
|--------------------|-------------------|--------------------|-------------------------|--------------------|
| | Aural Prompt (A) | Written prompt (W) | Aural prompt (A) | Written prompt (W) |
| N=40 | 10 | 10 | 10 | 10 |

4.2 Materials

a. *Aural and written prompts:* The aural and the written prompts were the aural and the written stimuli. That is, the test candidates wrote about the topic based on what they had understood from the prompt. The content of

both types of prompts was exactly the same (see Appendix 1). However, the way the prompts were presented was totally different. The written prompt was provided on papers, without the candidates being allowed to keep them during their writing time while the aural prompt was presented orally to the candidates two times without any interruption.

- b. *Concept mapping instruction text:*** Before the candidates in pre-task planning groups started their writing tests, they learnt concept mapping strategy by a short text as a sample prompt and its attached concept map (see Appendix 2). The papers were given to each individual and the test takers were given explicit explanations on how to make concept maps. They were required to extrapolate from this example to prepare a concept map and write their answer in response to a different prompt. They were allowed to practice the strategy for five minutes.
- c. *Pre test:*** A test was developed to ensure that the instructions and the allotted time for different planning conditions were suitable. Thus the writing test through pre-task planning condition and no planning condition was piloted with eight intermediate learners in the same language institute.

4.3 Assessment conditions and procedures

In this study, there are four groups:

- No planning-aural prompt (NOPA)
- No planning-written prompt (NOPW)
- Pre-task planning with aural prompt (PTPA)
- Pre-task planning with written prompt (PTPW)

Regardless of the type of prompt, in the pre-task planning condition the candidates were asked to apply the strategy of concept mapping in their planning time. But in the no planning condition, no extra time was provided for planning and consequently for the application of the strategy of concept mapping:

- a. *Pre-task planning with implementing the strategy of concept mapping:*** In this condition, the test takers did not start writing at once when they received the prompt (i.e. aural or written). But for five minutes, they planned their writings by seeking help from concept mapping strategy. Then, they had fifteen minutes to write at least 200 words about the given topic according to their planning maps. The provided planning time was based on the pilot test.

b. No planning without implementing the strategy of concept mapping: In the no planning condition, the test takers performed their writing tasks at once when they received the prompt (i.e. aural or written). They did not have any extra time for pre-task planning but had fifteen minutes and were expected to write at least 200 words about the given topic. The provided writing time was based on the pilot test and also the study by Ellis and Yuan (2004) in that they allowed seventeen minutes to the participants to write at least 200 words and develop their ideas under online pressured planning condition.

4.4 Measurement of variables

Enhancing complexity, accuracy, and fluency of produced language is a favorite goal in the world of EFL learning and teaching (Skehan, 1996). Therefore, the dependent variables of this study were considered to be the discourse measures of complexity, accuracy, and fluency of writing production. Also, each participant's written work was divided into T units and clauses. A T-unit is a main clause with all subordinate clauses embedded in it (Storch, 2005).

a. Complexity: Complexity of one's L2 system can be interpreted as its variety, and elaborateness (Housen & Kuiken, 2009). Biber, Gray and Poonpon (2011) revealed that writing complexity, which can be better captured by some phrasal rather than clausal features, is fundamentally different from oral complexity. Therefore according to their study, three non clausal complexity measures: Prepositional phrases functioning as noun modifiers (PP), attributive adjectives (AA), and nouns as nominal pre-modifiers per text (NP) were utilized.

Since Foster and Skehan (1996) consider proportion of clauses to T-units as a reliable measure which correlates well with other measures of complexity, it was also added. Although clausal and phrasal measures seem different and may not correlate well, both were used in this study in order to look at the same problem from different perspectives.

b. Accuracy: How much the rules of the target language are met refers to accuracy (Skehan, 1996). In this research, accuracy was measured by global units in terms of the proportion of error-free T-units to all T-units (EFT/T) and error-free clauses to all clauses (EFC/C). Both proportions were expressed by percentages. These global units correlate and represent a realistic measure of accuracy (Wigglesworth & Storch, 2009).

c. Fluency: When language is produced in real time without unreasonable pause or hesitation, it is fluent (Ellis & Barkhuizen, 2005). In other words, communicating language in real time makes it fluent (Skehan, 1996). However, it is not possible to measure the length of pauses in writing. Therefore, Wigglesworth and Storch (2009) pioneered a method to measure the fluency of writing production. Following them, in this study fluency was measured in terms of the average number of words per text (WPT), T-units per text (TPT) and clauses per text (CPT) (Rezazadeh, 2011).

It seems necessary to mention that Inter-rater agreement on different measures of complexity, accuracy, and fluency was done by Wigglesworth and Storch (2005) with a random sample of 12 texts which were being coded by a second rater. Inter-rater consistency for T-unit, clause and error free clause identification was 98%, 88% and 84% respectively. The researchers admitted that ensuring a high level of inter-rater reliability on accuracy was difficult to some extent.

5. Results

a. The effect of concept mapping strategy through aural and written prompts on test takers' writing production in terms of complexity, accuracy and fluency:

Since concept mapping strategy had only been used in pre-task planning groups, the two-way ANOVA was carried out on each dependent variable to determine for which measures the differences were significant in pre-task planning condition (i.e. pre-task planning with aural prompt and pre-task planning with written prompt).

Table 2. Descriptive statistics

| Planning Condition | Measure | Dependent variable | Prompt | Mean | Std. Deviation | N |
|--------------------|----------|--------------------|--------|----------|----------------|----|
| PTP | Fluency | TPT | A | 7.1000 | 1.52388 | 10 |
| | | | W | 7.9000 | 1.19722 | 10 |
| | | CPT | A | 16.8000 | 6.39097 | 10 |
| | | | W | 18.7000 | 3.97352 | 10 |
| | | WPT | A | 97.5000 | 24.71279 | 10 |
| | accuracy | (EFT/T)% | W | 1.0230E2 | 7.18099 | 10 |
| | | | A | 18.2830 | 14.90072 | 10 |
| | | | W | 22.6900 | 17.55863 | 10 |

| Planning Condition | Measure | Dependent variable | Prompt | Mean | Std. Deviation | N |
|--------------------|------------|--------------------|--------|---------|----------------|----|
| | | (EFC/C)% | A | 42.4500 | 19.67854 | 10 |
| | | | W | 35.8700 | 4.01775 | 10 |
| | | AA | A | 1.8000 | .78881 | 10 |
| | | | W | 3.0000 | 2.44949 | 10 |
| | complexity | NP | A | 1.8000 | 1.31656 | 10 |
| | | | W | 1.2000 | 1.39841 | 10 |
| | | PP | A | 3.3000 | 1.41814 | 10 |
| | | | W | 3.1000 | 2.46982 | 10 |
| | | C/T | A | 2.3161 | .46026 | 10 |
| | | | W | 2.3815 | .43281 | 10 |

Table 3. Summary of findings from two-way ANOVAs on fluency, accuracy, and complexity across prompt

| Source | Measure | Dependent variable | alpha | df | Mean Square | F | Sig. | Partial Eta Squared |
|--------|------------|--------------------|-------|----|-------------|-------|------|---------------------|
| | Fluency | TPT | .05 | 1 | 8.100 | 1.984 | .168 | .052 |
| | | CPT | .05 | 1 | 81.225 | 3.359 | .075 | .085 |
| | | WPT | .01 | 1 | 1822.50 | 4.686 | .037 | .115 |
| Prompt | Accuracy | (EFT/T)% | .01 | 1 | 912.694 | 2.396 | .130 | .062 |
| | | (EFC/C)% | .01 | 1 | 10.302 | .026 | .873 | .001 |
| | complexity | AA | .01 | 1 | 24.012 | 5.885 | .190 | .143 |
| | | NP | .01 | 1 | 8.100 | 3.025 | .091 | .078 |
| | | PP | .05 | 1 | .900 | .200 | .658 | .006 |
| | | C/T | .05 | 1 | .196 | .665 | .420 | .018 |

The results of two-way ANOVAs, illustrated in Table 3, show that there is no significant difference between application of the strategy of concept mapping through aural prompt and its application through written prompt in pre-task planning condition ($p > 0.01$ and $p > 0.05$). In order to satisfy the principal of homogeneity of variances for groups, alpha is considered 0.01 for some dependent variables and 0.05 for some others.

b. The effect of different planning conditions, no planning and pre-task, through aural and written prompts on test takers' writing production in terms of complexity, accuracy, and fluency:

The second research question aimed to investigate the interaction between the planning conditions and the prompts regarding the complexity, accuracy, and fluency of EFL learners' writing test performance. Therefore, two-way ANOVAs were carried out on each dependent variable in order to determine whether the differences were significant in terms of the discourse measures of complexity, accuracy, and fluency.

Table 4. Descriptive statistics

| Measure | Dependent variable | Planning Condition *Prompt | Mean | SD | N |
|----------|--------------------|----------------------------|----------|----------|----|
| Fluency | TPT | NOPA | 7.8000 | 1.98886 | 10 |
| | | NOPW | 8.8000 | 2.93636 | 10 |
| | | PTPA | 7.1000 | 1.52388 | 10 |
| | | PTPW | 7.9000 | 1.19722 | 10 |
| | CPT | NOPA | 18.6000 | 4.37671 | 10 |
| | | NOPW | 22.4000 | 4.57530 | 10 |
| | | PTPA | 16.8000 | 6.39097 | 10 |
| | | PTPW | 18.7000 | 3.97352 | 10 |
| | WPT | NOPA | 1.0560E2 | 16.41950 | 10 |
| | | NOPW | 1.2780E2 | 24.97910 | 10 |
| | | PTPA | 97.5000 | 24.71279 | 10 |
| | | PTPW | 1.0230E2 | 7.18099 | 10 |
| Accuracy | (EFT/T)% | NOPA | 10.0200 | 8.96447 | 10 |
| | | NOPW | 24.7200 | 30.21176 | 10 |
| | | PTPA | 18.2830 | 14.90072 | 10 |
| | | PTPW | 22.6900 | 17.55863 | 10 |
| | (EFC/C)% | NOPA | 37.7700 | 21.31812 | 10 |
| | | NOPW | 46.3800 | 27.29061 | 10 |
| | | PTPA | 42.4500 | 19.67854 | 10 |
| | | PTPW | 35.8700 | 4.01775 | 10 |
| AA | NOPA | 2.5000 | 1.77951 | 10 | |
| | NOPW | 4.4000 | 2.50333 | 10 | |
| | PTPA | 1.8000 | .78881 | 10 | |
| | PTPW | 3.0000 | 2.44949 | 10 | |

| Measure | Dependent variable | Planning Condition *Prompt | Mean | SD | N |
|------------|--------------------|----------------------------|--------|---------|----|
| complexity | NP | NOPA | .8000 | .78881 | 10 |
| | | NOPW | 3.2000 | 2.52982 | 10 |
| | | PTPA | 1.8000 | 1.31656 | 10 |
| | | PTPW | 1.2000 | 1.39841 | 10 |
| | PP | NOPA | 4.2000 | 1.68655 | 10 |
| | | NOPW | 3.8000 | 2.65832 | 10 |
| | | PTPA | 3.3000 | 1.41814 | 10 |
| | | PTPW | 3.1000 | 2.46982 | 10 |
| | C/T | NOPA | 2.4613 | .66369 | 10 |
| | | NOPW | 2.6755 | .58018 | 10 |
| | | PTPA | 2.3161 | .46026 | 10 |
| | | PTPW | 2.3815 | .43281 | 10 |

Table 5. Summary of findings from two-way ANOVAs on fluency, accuracy, and complexity across planning condition and prompt

| Source | Measure | Dependent variable | alpha | df | Mean Square | F | Sig. | Partial Eta Squared |
|--------------------|------------|--------------------|-------|----|-------------|-------|------|---------------------|
| Condition * Prompt | Fluency | TPT | .05 | 1 | .100 | 0.024 | .877 | .001 |
| | | CPT | .05 | 1 | 9.0125 | 0.373 | .545 | .010 |
| | | WPT | .01 | 1 | 756.900 | 1.946 | .172 | .051 |
| | Accuracy | (EFT/T)% | .01 | 1 | 264.865 | 0.695 | .410 | .019 |
| | | (EFC/C)% | .01 | 1 | 576.840 | 1.440 | .238 | .038 |
| | complexity | AA | .01 | 1 | 1.225 | 0.305 | .584 | .008 |
| | | NP | .01 | 1 | 22.500 | 8.402 | .006 | .189 |
| | | PP | .05 | 1 | 0.100 | .022 | .882 | .001 |
| | | C/T | .05 | 1 | .055 | .188 | .667 | .005 |

The results of two-way ANOVAs show that there is a statistically significant difference ($p \leq 0.01$) between the no planning group with written prompt and the pre-task planning group with written prompt in terms of nominal pre-modifiers per text with the no planning groups' better performance. There was another significant difference ($p \leq 0.01$) between the no planning group with aural prompt and the no planning group with

written prompt in terms of nominal pre-modifiers per text. The no planning group with written prompt performed better than the no planning group with aural prompt.

6. Discussion and Conclusion

This study was aimed at examining the effects of aural and written prompts through no planning and pre-task planning conditions on EFL learners' writing test performance in terms of complexity, accuracy, and fluency. The results illustrate that planning generally proves counterproductive in assessment setting. This is specifically obvious in the case of complexity and fluency. The results also show that aural or written prompt makes no changes in the situation. Even the strategy of concept mapping, applied during the pre-task planning time, is not beneficial to complexity, accuracy, and fluency of the written texts. In this section, the findings of the study as well as how they agree with those of the previous studies will be discussed.

a. The effect of concept mapping strategy through aural and written prompts on test takers' writing production in terms of complexity, accuracy and fluency

The results revealed that there was no significant difference between the quality of the written texts produced by concept mapping through aural prompt and the quality of the ones produced by concept mapping through written prompt under pre-task planning condition in assessment setting. In other words, the type of the received prompt, aural or written, did not make any significant difference in the way the test takers benefited from the strategy of concept mapping.

The prompts provided some information about the topic that the test takers wrote about. The way the prompts fed the test takers' minds was very similar. Actually, both types of map processing instigators worked in the same way.

The results revealed that the test takers who received the written prompt in the pre-task planning condition utilized their time better to produce more fluent written texts (with regard to T-units per text, clauses per text, and words per text) than those who performed in the same planning condition with aural prompt. Although, the difference was not significant, the pre-task planning with the written prompt had a good effect on the number of words per text as an aspect of the fluency measure. In this case, the written prompt may have formed a stronger base for the application of

concept mapping strategy and has supported better organization of the ideas which is an important role of concept mapping strategy. Possibly, the written prompt has acted as a visual aid for concept mapping, in which the regularity in events and objects are graphically demonstrated (Novak, 1992). The regular pattern the maps follow can be strongly relevant to the smooth flow of the produced language because they can create images of the words in the mind of the individual who will map the similar concepts later. Written prompt can lead to better recall of the concepts than the aural prompt that creates the images mentally.

The concept mapping developers emphasize the assimilation of new information into the students' prior knowledge (Novak, 2010). Thus, the foreknowledge is awakened and combined with new information (Beidogan & Bayindir, 2010). Written prompt is probably better involved in this building process. It helps the test taker to assimilate new ideas to their prior knowledge and visually feeds the created maps by faster building of an image of the concepts. It is probably the reason which accounts for the difference in the fluency outcomes, although the differences do not prove to be significant.

In the case of accuracy, mixed results have been obtained: For the proportion of error free T-units of all T-units, written prompt works better than aural prompt. But for the proportion of error free clauses of all clauses, aural prompt is more useful. Since for none of the accuracy measures the differences are significant between the written texts of individuals in the pre-task planning condition with written prompt and those in the pre-task planning condition with aural prompt, these mean differences are not taken into account. According to the results, aural and written prompts have very similar roles in establishing the concept maps which ease the synthesis of information (Novak & Gowin, 1984), leading to better monitoring of the outcome as an inevitable characteristic of a more accurate language production (Ellis, 2005). In other words, aural and written prompt are similarly involved in supporting the map buildings and consequently in more accurate utilization of the drawn maps. Accordingly, different types of prompts are only map instigators without any special influence on the quality of the produced text.

Coming to the complexity measures, the results revealed that there is no significant difference between the written test performance of the individuals in pre-task planning condition with aural prompt and their performance in pre-task planning condition with written prompt with regard

to the lexical complexity measures (i.e. attributive adjectives per text, nominal pre-modifiers per text, and post modifying prepositional phrases per text) and the grammatical complexity measure (proportion of clauses to T-units). Actually, making a conceptual map pattern which makes thought and meaning (Pieronek, 1994) was affected by both prompts in the same way.

When the individuals apply concept mapping strategy, they add some concepts to the hierarchies of their maps (Gul & Boman, 2006) which may be a set of lexical or grammatical features that make the written complexity. When the test takers receive aural prompt, they start making some maps with the concepts and ideas motivated by the prompt. The candidates who receive written prompt also try to do the same task. This study shows that the number of added concepts and the grammatical features per text were very slightly different between the two groups. Written prompt closely resembled aural prompt in feeding the created concept maps that can inject the suitable concepts for making complex written language. The prompts similarly provoke the test takers to make and add different concepts to their written text.

b. The effect of different planning conditions with aural and written prompts on test takers' writing production in terms of complexity, accuracy, and fluency

The findings of the study suggest that no planning condition with written prompt has a significant effect on the measure of nominal pre-modifiers per text as a lexical aspect of written complexity in comparison with pre-task planning condition with written prompt. The findings also reveal that the interaction of no planning and pre-task planning with aural or written prompt did not have any advantage to written complexity in terms of attributive adjectives per text, post modifying prepositional phrases per text, and proportion of clauses to T-units.

This study tried the new measures of written complexity which are phrasal (Biber et al., 2011) in contrast to the previous studies in which the complexity measures are clausal. Therefore, the results of the current study cannot be regarded as to be in line with any of the previous planning studies in terms of the new complexity measures. Since no planning condition with written prompt was found to be more effective in terms of nominal pre-modifiers per text than pre-task planning with written prompt, pre-task planning proved to be counterproductive in testing context in terms of one aspect of written complexity measure.

Considering the mentioned grammatical complexity (i.e. proportion of clauses to T-units), the interaction of no planning and pre-task planning with aural or written prompt was not advantageous to the quality of the written text. It is in line with the abovementioned investigations on the effect of planning time on test performance in which planning time may not be associated with more complex language production. A possible explanation for the results regarding complexity is the human's limited attentional resources; that is the students have to give more importance to form or to meaning during language production (Skehan, 2009). Although planning can make up for the limitations (Yuan & Ellis, 2003), test takers may not benefit from planning time because of giving considerable attention to avoiding errors (Wigglesworth & Elder, 2010). Therefore, the concept maps created during pre-task planning time may have not helped the test takers' formulation and execution of their written texts (see Kellog, 1996).

In the case of fluency and accuracy, the findings are consistent with those of Wigglesworth (1997), Wigglesworth (2000), Iwashia et al., (2001), and Wigglesworth and Elder (2010). Their findings on the effect of planning time on test performance gave tentative support to the hypothesis that planning time may be associated with more complex, more accurate and more fluent language production. Although the present study tried to change the testing situation by adding different prompts before the planning chance, no better results in terms of fluency and accuracy were obtained. In other words, no-planning condition with aural prompt or with written prompt and pre-task planning condition with aural prompt or with written prompt did not significantly affect the fluency and the accuracy measures.

The possible reason for the obtained results is that in testing context the test takers feel pressured. Therefore, their planning may be hurried. Since students' comfort with map making is a crucial factor behind proper application of the strategy, under the pressure of assessment setting test takers may find their maps confusing. The maps may be simply some barriers to their flow of language because they may not be able to choose 'the right words or phrases' when they want to connect the concepts (Gul & Buman, 2006, p.204). Possibly, the planning chance does not prove to be useful although it accompanies aural or written prompt and does not lead to better fluency and accuracy of the written texts.

The findings of the study also suggest that no planning condition with written prompt has a significant effect on the measure of nominal pre-modifiers per text as a lexical aspect of written complexity in comparison

with the same planning condition with aural prompt. Maybe written prompt has put less pressure on test takers' minds when they have wanted to recall the prompt content in order to write about. In the atmosphere of testing context, reading a text can provoke less anxiety than listening to it. Because of the less pressure the test takers feel, they may be able to invest their energy in creative and complex language use in terms of nominal pre-modifiers per text.

By and large, it is concluded that the test takers' writing performance has been slightly affected by the planned conditions, concept mapping strategy and the received prompts with regard to complexity, accuracy, and fluency measures.

7. Implications

a. Written and aural prompts as concept mapping stimuli in writing test performance:

Written and aural prompts can act as schema builders. They help the learners to develop their ideas based on the received prompt. Both prompts can be applied before classroom writing tasks or writing tests. They work as bases of what the students will develop later in their writings. The results of the current study reveal that the prompts both feed the test takers' minds very similarly. They act as map processing instigators in the same way, leading to similarly complex and accurate written texts.

When fluency of writing is important to the teacher, s/he can create a written prompt to help her/his students reach this purpose. The written prompt can support the better organization of the ideas and work as a visual aid for the mapping process which evolves the graphical flow of events and ideas. This possibly will manage easy flow of language.

b. Written prompt as a stimulus in writing test performance:

When the online planning is pressured (i.e. no planning time is provided) in testing context, a written prompt can be a very helpful stimulus to motivate the students to write and develop their ideas, using a complex language with regard to the nominal pre-modifiers per text as a written complexity measure.

c. No planning in writing test performance:

No planning in testing context can be more fruitful than pre-task planning since the former can make the test takers use their language resources to

produce more complex and fluent written texts. Although pre-task planning condition has a slight positive effect on written accuracy, this small effect is too small to be paid attention to.

8. Limitations

A number of limitations need to be acknowledged with regard to the interpretation of the results:

First, the sample size was small. Therefore, the insignificant results could be assumed to be the consequence of the small sample not the small effect of the independent variable. Although the sample size can be considered a limitation, it does not cause any problem in assessing the results of the current study. The reason is that measures of effect size in ANOVA are measures of the degree of association between the effect of the independent variable and the dependent variable. If the value of the measure of association is squared it can be interpreted as the proportion of variance in the dependent variable that is attributable to each effect. Partial Eta squared is estimate of the degree of association for the sample which is displayed by SPSS when you check the display effect size option. Pearson's correlation is widely used as an effect size when paired quantitative data are available. Pearson's r can vary in magnitude from -1 to 1 , with -1 indicating a perfect negative linear relation, 1 indicating a perfect positive linear relation, and 0 indicating no linear relation between two variables. Specifically, the effect size is estimated as small, medium and large when r is 0.1 , 0.3 , and 0.5 (Cohen, 1988). In this study, apart from the sample size, the influence of the independent variables was very little.

Second, the instruction of concept mapping strategy occurred in a short time because extra sessions were impossible. Actually, the test takers did not have enough time to get very familiar with the strategy.

The third limitation to this study is how the results are interpreted with regard to pre-task planning time and concept mapping strategy applied in the planning time. It is not possible to measure to what percent the strategy and to what percent pre-task planning is responsible for the obtained results.

The fourth, two different types of measures have been utilized with regard to the written complexity in the current study. But it is not clear what type measures the written complexity to a greater extent.

Finally, the sampling was non-random. The participants were selected based on their accessibility and their willingness to volunteer. In relation to the experiments conducted under natural classroom context Ellis (2011)

argues that although there are problems with conducting a study in intact classes, the random assignment of treatment to these classes makes the study quasi-experimental. On the other hand, Ary, Cheser Jacobs, Razavieh and Sorensen (2010, p. 296) argue that "an experiment conducted under a more natural environment such as a classroom may have greater external validity, but its internal validity may be less". They state that when an experiment is strictly controlled it is internally valid but the increased artificiality makes it less externally valid and less generalizable. Based on the above quotation, we can argue that since this research was conducted under a natural classroom situation, though less internally valid, its external validity is high.

References

- Ausubel, D. (1963). Cognitive structure and the facilitation of meaningful verbal learning. *Journal of teacher education*, 14, 217-221.
- Ahmadian, M.J., & Tavakoli, M. (2010). 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.
- Ary, D., Cheser Jacobs, L., Razavieh, A., & Sorensen, C. (Ed.). (2010). *Introduction to research in education*. Wadsworth: Cengage Learning.
- Beidogan, H.O., & Bayindir, G. (2010). Effect of concept map supported teaching approaches from rules to sample and sample to rules to grammar teaching. *Procedia Social and Behavioral Sciences*, 2, 3954-3964.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should We Use Characteristics of Conversation to Measure Grammatical Complexity in L2 Writing Development? *TESOL QUARTERLY*, 45 (1), 5-35.
- Chularut, P., & DeBacker, T.K. (2003). The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language. *Contemporary Educational Psychology*, 29, 248-263.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Elder, C., & Iwashita, N. (2005). Planning for test performance: What difference does it make? In R. Ellis (Ed.). *Planning and task performance in a second language* (pp. 219-238). Amsterdam: John Benjamins.

- Elder, C. A., Iwashita, N., & McNamara, T. (2002). Estimating the difficulty of oral proficiency tasks: What does the test-taker have to offer? *Language Testing, 19*, 347–368.
- 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. (2009). The Differential Effects of Three Types of Task Planning on the Fluency, Complexity, and Accuracy in L2 Oral Production. *Applied Linguistics, 30*(4), 474–509.
- Ellis, R. (2011). Editorial. *Language Teaching Research, 15* (2), 133-135.
- 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.
- Ellis, R. & Barkhuizen, G. (2005). *Analyzing Learner Language*. Oxford University Press.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 26* (1), 59-84.
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research, 3*, 215-247.
- Gul, R.B., & Boman, J. (2006). Concept mapping: A strategy for teaching and evaluation in nursing education. *Nurse Education in Practice, 6*, 199–206.
- Housen, A., & Kuiken, F. (2009). Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics, 30*(4), 461–473.
- Iwashita, N., Mc Namara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an Information-Processing Approach to Task Design. *Language Learning, 51*(3), 401-436.
- Kellogg, R. (1996). A model of working memory in writing. In C. Levy and S. Ransdell (Eds.). *The Science of Writing* (pp. 57–73). Mahwah NJ: Lawrence Erlbaum.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge MA: The MIT Press.
- Liu, L.P. (2011). A study on the use of computerized concept mapping to assist ESL learners' writing. *Computers & Education, 57*, 2548–2558.

- Novak, D.J. (1992). *The Theory Underlying Concept Maps and How to Construct Them*. Cornell University. Retrieved April 20, 2004, from <http://cmap.coginst.uwf.edu/info/>
- Novak, D.J. (2010). Learning, Creating, and Using Knowledge: Concept maps as facilitative tools in schools and corporations. *Journal of e-Learning and Knowledge Society*, 6(3), 21-30.
- Novak, J.D., & Gowin, D.B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Ojima, M. (2006). Concept mapping as pre-task planning: A case study of three Japanese ESL writers, *System*, 34, 566–585.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109–148.
- Ortega, L. (2005). What do learners plan? Learner-drive attention to form during pre-tasks planning. In R. Ellis (Ed.). *Planning and task performance in a second language* (pp. 77–109). Amsterdam: John Benjamins.
- Rezazadeh, M. (2011). *The Effects of Individual and Collaborative Planning on Fluency, Complexity, and Accuracy in L2 Written Production of Iranian EFL Learners: Focusing on the Relationship among Planned Conditions, Language Proficiency, and Task Type*. Unpublished master's thesis, University of Isfahan, Isfahan, Iran.
- Pieronek, F. (1994). Using maps to teach note taking and outlining for report writing. *The Social Studies*, 85, 165–169.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57.
- Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.). *Planning and task performance in a second language* (pp. 111–141). Amsterdam: John Benjamins.
- Schultz, M. (1991). Mapping and cognitive development in the teaching of foreign language writing. *The French Review*, 64, 978–988.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.
- Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics*, 18, 268–286.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. Oxford University Press. *Applied Linguistics*, 30(4), 510–532.
- Skehan, P. & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1, 185–211.
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14, 153–173.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Eds.). *Planning and task performance in a second language* (pp. 239-273). Amsterdam: John Benjamins.
- Wendel, J. N. (1997). *Planning and second language narrative production*. Unpublished doctoral dissertation, Temple University Japan Campus, Japan.
- Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14, 85-106.
- Wigglesworth, G. (2000). Issues in the development of oral tasks for competency-based assessments of second language performance. In G. Brindley (Eds.). *Studies in immigrant English language assessment*. Sydney. National Centre for English Language Teaching and Research Macquarie University.
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466.
- Wigglesworth, G., & Elder, C. (2010). An investigation of the effectiveness and validity of planning time in speaking test tasks. *Language Assessment Quarterly*, 7(1), 1-24.
- Yuan, F. & Ellis, R. (2003). The effects of pre task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1-27.

Appendices

Appendix1: The prompt text

Do you agree or disagree with the following statement? Most experiences in our lives that seemed difficult at the time become valuable lessons for the future. Use reasons and specific examples to support your answer.

Appendix2: Concept mapping instruction text

Difficult experiences are the best teachers. I remember the first time I had to give a presentation to my classmates. I was very shy and afraid to speak in front of the whole class. I spent a long time preparing for my presentation. When I gave my presentation, everyone listened. They asked questions and I could answer them. Now I know I can talk in front of the class and do a good job.

