ADIYAMAN UNIVERSITY
Journal of Educational Sciences
(AUJES)

https://dergipark.org.tr/tr/pub/adyuebd

Effects of Task Complexity on Text Easibility and Coherence of EFL Learners' Narrative Writing

Mine Yıldız[1], Savaş Yeşilyurt[2]
[1]Atatürk University, Department of Foreign Languages Education, Erzurum, Turkey (iD)
[2]Atatürk University Department of Foreign Languages Education, Erzurum, Turkey (iD)

Effects of Task Complexity on Text Easibility and Coherence of EFL Learners' Narrative Writing

Mine Yıldız[1*], Savaş Yeşilyurt[2]
[1] Atatürk University Department of Foreign Languages Education, Erzurum, Turkey
[2] Atatürk University Department of Foreign Languages Education, Erzurum, Turkey

Abstract

This study was carried out to examine the effects of task complexity on text easibility and coherence in narrative writing of EFL learners. Data were collected from 41 Turkish EFL learners during a writing course. Task complexity was operationalized at two levels as a complex and simple task based on the resource-dispersing variables of Robinson's the Triadic Componential Framework, +/- task structure. Accordingly, a colorful picture was first illustrated on the board, and students were asked to examine the picture for five minutes (complex task /-TS). They were then asked to write a story based on the picture they had seen (simple task /+TS). Two weeks later, they were given a sheet involving 16 pictures designed in an order and asked to narrate a story based on these pictures. Their essays were analyzed by the researcher and another rater in terms of coherence through an analytic rubric. An automated program was used to evaluate the essays for text easibility indices involving the indices of narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion. The results analyzed with a Wilcoxon signed-rank test showed that complexity of the task had a statistically significant effect on some indices of text easibility such as word concreteness, and referential cohesion, whereas other indices, narrativity, syntactic simplicity, and deep cohesion, and coherence in their writing production were not affected by the complexity of the task at a significant level. However, it can be concluded that students' texts produced in simple tasks are easier to comprehend.

Keywords: Task-based language teaching, Task complexity, Coherence, Text-easibility, Coh-Metrix

Introduction

Task-based language teaching (TBLT) aims to enhance language learning based on the tasks involving meaningful, pragmatic, and communicative activities in all processes of learning such as planning, instruction, and assessment (Larsen-Freeman & Anderson, 2011). TBLT, one of the examples of a 'strong version' of the communicative approach, has received great attention of not only second language acquisition researchers but also researchers of second language teaching as it is primarily motivated by a theory of learning (Richards & Rodgers, 2001) and poses several advantages over PPP (present-practice-produce) paradigm that is claimed to be an over-simplified approach (Kırkgöz, 2014). In other words, the learners in TBLT are provided with the opportunity to learn the language through authentic scenarios involving meaningful, intentional, pragmatic, and surely communicative activities in which they are required to use their linguistic resources to perform the task. (Arslanyilmaz, 2013; Willis, 1996). Hence, they are able to use the language as a "vehicle for attending task goals' (Willis, 1996, p. 25) in a meaningful and natural atmosphere inside the classroom.

The main focus of a task is on meaning and the primary aim is to achieve an outcome as a result of the process using language (Skehan, 1998). Similarly, although Ellis (2009) states that "there is no single way of doing TBLT" (p. 224), he proposes three phases involving the pre-task phase, the main phase (the only one to be obligatory), and the post-task phase. Willis (1996), likewise, presents three elements but shows a difference in their features as pre-task, task cycle, and language focus (Willis, 1996, p. 135). Accordingly, a topic or task is introduced at the pre-task phase through clear and insightful instructions; the learners are required to conduct, plan how to report the outcome of their performance, and then to produce something at the phase of task cycle, and lastly, they analyze the recording of their reports and practice the phrases, words or structures at the phase of language focus. Therefore, writing activities can be described as a task since they meet all the requirements of a task. Their focus is on meaning, they involve a goal to be achieved by learners, and the learners obtain an outcome as a result of their performance. Furthermore, the writing activities are conducted at three phases as pre-writing, while-writing, and post-writing. Thus, students who participated in this study were asked to produce narrative essays for their task performance.

---
[*] Corresponding Author: *Mine Yılıdız, mine.yazici@atauni.edu.tr*

Studies on Task Complexity

The studies reviewed in this study mostly are based on the two influential frameworks, Robinson's Cognition Hypothesis (Robinson, 2001, 2003, 2005; Robinson & Gilabert, 2007) and Skehan's Limited Attentional Capacity Model (Skehan & Foster, 1999, 2001), in terms of either dependent or independent variables, or both of them (Ellis & Yuan, 2004; Jackson & Suethanapornkul, 2013; Kawauchi, 2005; Kim, 2020a, 2020b; Ruiz-Funes, 2015; Tavakoli & Skehan, 2005; Yang, 2014; Yang, Lu, & Weigle, 2015; Yıldız & Yeşilyurt, 2017; Yuan & Ellis, 2003). Moreover, task complexity is generally identified along with the variables of these two competing models. Limited Attentional Capacity Model (Skehan, 2003, 2014; Skehan & Foster, 1999, 2001), which proposes that while performing a complex task people use more attentional resources due to a limited capacity they have in order to process information, deals with the task complexity under three dimensions such as code complexity, cognitive complexity, and communicative stress. While the focus of the code complexity is on linguistic demands the task requires, the content of the task and structuring of the material used in the task is seen within the concern of the cognitive complexity which is also divided into two main categories as *cognitive familiarity* and *cognitive processing*. The third dimension, communicative stress, basically focuses on the conditions and components of task performance such as participants, presentation, text, and time. Furthermore, it suggests that manipulations regarding the task require the learners to use more cognitive demand and attentional resources and thus reach production with trade-off effects among the three basic constructs complexity, accuracy, and fluency. Due to the limited attentional capacity learners have, they are not capable of paying simultaneous attention to those dimensions of language; that is, while paying attention to one dimension, they fail to focus on the others. In other words, focusing on producing complex performance probably leads to trade-off effect between accuracy and fluency since over-attention to complexity probably leads to lacks in accuracy or fluency of the production, or vice versa (Skehan, 2009).

The other model on which task complexity studies are based, Robinson's Triadic Componential Framework, similarly deals with the task complexity under three dimensions: task complexity, task conditions and task difficulty. This framework considers task complexity within the frameworks of information-processing demands required for a pedagogic task for memory, attention, and reasoning (Robinson, 2001). It presents two categories for the cognitive task features as resource-directing and resource-depleting variables, which was seen with a new name in the expanded version of Cognition Hypothesis as resource-dispersing variables (Robinson & Gilabert, 2007). The tasks manipulated along the variables of this framework are assumed to have language production in various ways. Whether there are few elements to be compared (+/- few elements) or not, whether the events occur in the past or present, or things are far or near (+/- here-and-now), whether some reasoning demands are provided for the learner (+/- reasoning) or not is assumed to require *cognitive* and *conceptual* demands and involved in resource-directing variables. On the other hand, the dimension of resource-dispersing variables deals with other variables such as *performative* and *procedural* demands for learners manipulated along whether planning time is allocated to the learners or not (+/- planning), the task has a loose or tight structure (+/- single task), and learners have prior knowledge or not (+/- prior knowledge) in order to apply in their task performance (Robinson, 2001). Furthermore, in contrast to Skehan (1998) suggesting that learners have to prioritize between the three dimensions due to their capacity and attentional resources to process information, Robinson (2001) suggests that learners' performance can be enhanced in all three dimensions of complexity, accuracy, and fluency (CAF).

The studies investigating whether increasing complexity of the task had any effect on language production show differences from many perspectives such as in terms of both dependent and independent variables although complexity, accuracy, and fluency are commonly preferred as dependent variables. For instance, Arslanyilmaz (2013) contrasted two different teaching tools – the computer-assisted task-based instruction (CATBI) and computer-assisted form-focused language instruction (CAFFI) in order to investigate the role of the teaching approach in second language development in terms of accuracy, lexical complexity, and fluency. According to the results of the study, the students taught through CATBI produced better language than those taught through CAFFI; in particular, although no significant difference for lexical complexity was seen, the language of production of task-based instruction was more fluent and accurate.

Tavakoli and Foster (2008) examined how oral second language performance is affected by narrative structure (tight/loose) and narrative complexity (with or without background information) in terms of the most common measures of task complexity, complexity, accuracy, and fluency (CAF). In support of previous studies, they concluded that accuracy appeared to increase through tight task structure and also that narrative tasks with background information seemed to result in higher syntactic complexity. In another study investigating the effects of task design on L2 task performance in terms of accuracy, fluency, syntactic complexity, and lexical diversity, Tavakoli (2009) pointed out that syntactic complexity could be enhanced through more structured tasks – narratives with both foreground and background storylines and also that they yielded more accurate and

fluent performance in more structured tasks than they did in less structured ones. However, no clear result has been obtained for the effect of task structure on lexical diversity.

Based on the cognitive frameworks for TBLT, Révész (2011) conducted a study with the goal of exploring whether there is a relationship between task complexity and learners' use of form-meaning mappings in oral tasks and also whether individual differences have an impact on such a relationship. Speech production of the participants performing two versions of the same argumentative task – complex or simple - manipulated along the +/- reasoning and the +/- few elements dimensions were analyzed through some global and specific measures of oral performance. It was illustrated that although participants' speech in the complex task was more accurate and lexically diverse but lower syntactically complex speech, no significant effects of learners' individual differences were observed.

In their study, Kuiken and Vedder (2007) firstly aimed to compare the two most influential models of task complexity – Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity Model – regarding the effect of task complexity on L2 writing performance in terms of three measures of linguistic complexity and accuracy. The learners of Italian and French were assigned two writing tasks manipulated along with cognitive complexity as a non-complex condition in which they were required to write a letter taking three requirements into consideration and a complex condition in which six requirements would be considered. Although in previous studies Kuiken et al., (2005), Kuiken and Vedder (2008) revealed an effect of task complexity on accuracy evaluating it through general measures, the study by Kuiken and Vedder (2007) utilized more specific measures of accuracy and lexical variation regarding error type and the most frequent words used to illustrate their role for such an effect. The results of the study confirmed that fewer errors were seen in the complex task which might explain the accuracy case in the complex task; in other words, the fact that complex tasks yield more accurate texts probably results from a decrease of lexical errors in such tasks. As for the frequency of words, while French participants used less frequent words in a complex task, the case for the Italian participants was the opposite. In light of the results, it was also pointed out that it seemed not possible to establish a relationship between task complexity and language proficiency level.

Operationalizing task difficulty as the storyline structure – loose or tight – Ahmadian et al., (2012) investigated the effect of task difficulty on self-repair behavior in L2 oral performance. While performing the structured task the participants were observed to mainly focus on producing error-free units in terms of lexicon, grammar, and phonology; on the contrary, in the unstructured task, they were primarily concerned with conceptualizing the oral production producing D- repairs (different information involving alteration of the content of the preverbal message) and A-repairs (appropriacy that includes changes in the content of the message in terms of inaccuracy, incoherence, ambiguity, and inappropriacy) regularly.

Similarly, Adams et al., (2015) investigated the role of task structure and language support in increasing the accuracy and linguistic complexity of writing via text chat. For their four experimental groups, they implemented two task variables – task structure (+/- TS) from Robinson's (2007) Triadic Componential framework and language support (+/- LS) utilizing pre-task to raise consciousness. Whereas learners in the +TS case were provided with detailed written instruction about task performance and also a worksheet guiding them, those in the condition of low task structure (-TS) were given just basic instructions but no worksheet. Similarly, learners of +LS condition were provided pre-task language support activities, but others did take no language support in their task that is therefore expected to be more complex. Analysis of the chat texts on the engineering simulation task revealed that although the learners performing more complex tasks (-TS and –LS) produced less accurate texts, making tasks more complex had no impact on the linguistic complexity.

In order to investigate whether cognitive task complexity influences lexical and syntactic complexity, Frear and Bitchener (2015) utilized resource-directing variables (Robinson, 2007) by manipulating the number of reasoning demands (+/- reasoning) and numbers of elements (+/- few elements). As a result of their analysis of letters by L2 writers of English in terms of lexical variety through a mean segmental type-token ratio and syntactic complexity by the ratio of dependent clauses to T-units, it was pointed out that an increase appeared on the lexical complexity as a result of increasing complexity of cognitive task. However, in contrast to the expectation of the Cognition Hypothesis (Robinson, 2001, 2007), in which it is assumed that increases in task complexity will lead to language development resulting in complex language performance, no significant change was seen in syntactic complexity among tasks.

Like many researchers based on the assumptions of Robinson's Cognition Hypothesis, Salimi et al., (2011) investigated the effects of tasks manipulated along with resource-directing factors on accuracy, fluency, and syntactic complexity. Using two versions of the same decision-making task, complex and simple, their findings on fluency and complexity confirmed the predictions of Cognition Hypothesis that complex tasks would lead to more fluent and syntactically complex texts; nevertheless, the case for accuracy was different. No significant difference was obtained between complex and simple tasks in accuracy.

The studies in literature scarcely investigate the writing task performances of learners in terms of task complexity (Jackson & Suethanapornkul, 2013). For instance, in their study reviewing the studies on task complexity, Salimi and Dadaspour (2012) revealed that many of the studies regarding the effects of task complexity mainly focus on L2 oral performance but just a few on the written performance of L2 learners. Furthermore, Ellis and Yuan (2004), Kormos (2011), Salimi et al. (2011) and Yang et al. (2015) drew attention to a limited number of studies on the effect of task complexity on L2 writing performance. Therefore, this study focuses on whether the narrative writing performance of EFL learners is affected by task complexity manipulated along with resource-dispersing variables (+/- task structure) by Robinson's Cognition Hypothesis. Furthermore, there is very limited research investigating the effects of task complexity on written production in the Turkish EFL context. The studies found in the Turkish context were the study of Genç (2012) that investigates the effects of strategic planning on the accuracy of EFL learners' both oral and written narrative task performances and that of Yıldız and Yeşilyurt (2017) examining whether task planning had an influence on the complexity and overall writing quality of EFL learners' writing performance. Therefore, this study aims at probing into the effects of cognitive task complexity on written production of EFL learners in Turkey.

Text Easibility And Coherence

The present study was carried out to see the effect of task complexity on L2 writing performance in terms of two dependent variables as text easibility and coherence. Therefore, the two most commonly interchangeably used terms, coherence, and cohesion should be explained in detail to show the distinction between these terms. Though being an important characteristic of effective writing in terms of connectedness that "refers to all of the links, both explicit and implicit, in a text that make it a unified whole" (Watson Todd et al., 2007), coherence is generally thought to be an abstract and fuzzy term to define exactly and make a distinction from other concepts in writing such as cohesion, unity, etc. Lee (2002) describes coherence as "the relationships that link the ideas in a text to create meaning for the readers" (p. 135). It is commonly misused with the term cohesion: whereas cohesion, in simple terms, regards implicit links, coherence refers to the opposite, explicit links (Watson Todd et al., 2007). In other words, whereas cohesion is described as the connection of ideas at sentence level or "the connectivity of ideas in discourse and sentences to one another in text, thus creating the flow of information in a unified way" (Hinkel, 2004, p. 279), being a broader term, coherence is the organization of ideas at discourse level with all elements.

Coherence is, in simple terms, what the reader grabs from the text while cohesion provides the reader with linguistic elements-cohesive devices- to make a connection between ideas (Crossley et al., 2016). As stated in their seminal work "Cohesion in English", regarded as a theoretical framework on textual cohesion, Halliday and Hassan (1976) describe cohesion as a semantic concept that illustrates "relations of meaning that exist within a text" (p. 4). Similarly, according to Harmer (2004), writers use two main elements to build cohesion in text-linguistic techniques and grammar structures; in other words, like Halliday and Hassan (1976), he also describes cohesion in two headings- lexical cohesion and grammatical cohesion. The coherence that enables the reader to catch both "the writer's purpose" and "the writer's line of thought" is far beyond the sentence level and achieved through sequencing information in order to meet the expectations of the discourse community that it is written for (Harmer, 2004, p. 22-25).

As in definitions of coherence and cohesion, the research shows also differences in the ways or measures to assess them. For example, one of the scales applied to assess coherence both in spoken and written discourse is a topic-based analysis which depends on identifying key terms in a text, finding the relationships between these terms, ranking these relationships, and then mapping the text along the hierarchy identified through the relationships (Todd et al, 2004). In their study, Todd et al. (2004) applied topic-based analysis because it meets the three criteria defined by the researchers to select an appropriate scale to evaluate coherence: it (1) is objective, (2) unequivocally measures coherence, and (3) focuses on propositional coherence that is predominant in written discourse rather than interactional coherence seen in informal spoken language. Similarly, Knoch (2007) reported that the previous scales developed to assess coherence are either too time-consuming or complicated. Therefore, in his study undertaken in three phases as (1) analysis of writing samples, (2) rating scale design, and (3) rating scale validation, he chose and adapted a topical structure analysis (TSA) scale with the aim of investigating whether the use of a TSA scale-an empirically-based scale- to evaluate coherence in the written production of students is more reliable and has greater discrimination compared to the more traditional measures. However, the results revealed that although raters using the TSA scale scored more accurately, the TSA scale was not less time-consuming than the previous scale; rather, it might require more labor to analyze a large number of written texts and thus not practical in some cases.

McNamara et al. (2009) used Coh-Metrix-an automated tool- to examine whether the quality of the essays- low or high- can be predicted through the three indices as syntactic complexity, lexical diversity, and word frequency. In contrast to the general notion that more cohesive and thus more coherent essays are produced by more proficient writers, their study using linguistic indices of cohesion from Coh-Metrix could not

provide any evidence about whether there is a significant difference between high- and low-proficiency essays in terms of coherence; that is, the essays scored high were not more coherent than those rated low (McNamara et al., 2009). According to McNamara et al. (2009), the Coh-Metrix cohesion indices validated by a number of studies are confidential to assess cohesion. In the light of the literature, they reached a conclusion that Coh-Metrix is "an extremely powerful text analysis tool, capable of assessing and differentiating an enormous variety of text types from the genre level to the sentence level" (p. 59). Therefore, in their study investigating the degree to which these indices have a role in predicting the quality of essays, they used 26 linguistic indices of cohesion from Coh-Metrix. Similarly, McNamara et al. (2010) point out that Coh-Metrix which is a tool presenting a great variety of linguistic indices for the automatic analysis of text comprehension uses lexicons, latent semantic analysis (LSA), and many other linguistic components and thus meets the needs of researchers who seek a computational linguistic analysis of texts to measure text cohesion and text difficulty in terms of various linguistic features such as word, sentence, paragraph, and discourse dimensions. Furthermore, their study comparing the outcomes of Coh-Metrix indices with two commonly used readability indices – Flesch Kincaid Grade Level and Flesch Reading Ease added evidence on validation of Coh-Metrix as a tool to assess cohesion.

Through the studies on cohesion, we reached the conclusion that we can measure text easibility indices through the automated tool "The Coh-Metrix Common Core Text Ease and Readability Assessor (T.E.R.A.)" which is less time-consuming and more practical. On the other hand, the coherence that is more subjective and exists in the mind of the reader will be best assessed using an analytical rubric that involves a specific dimension of coherence. This study which was carried out to see whether increasing complexity of writing task along with resource-directing variables as presence or absence (+/-) of few elements affect cohesion and coherence of EFL learners' narrative writing will probably contribute to filling the research gaps in the effects of task complexity, particularly in terms of written task performance. As also pointed out by Jackson and Suethanapornkul (2013) in their synthesis and meta-analysis study of research on task complexity, writing task performance has been rarely investigated in terms of task complexity although writing production of learners is valuable to obtain more reliable and concrete results. In line with these aims, the findings of the study were presented and discussed to provide responses to the following research questions:

1. Does increasing the complexity of a task affect the text easibility of EFL learners' narrative writing?

a. Does increasing the complexity of a task affect the text narrativity of EFL learners' narrative writing?

b. Does increasing the complexity of a task affect the syntactic simplicity of EFL learners' narrative writing?

c. Does increasing the complexity of a task affect word concreteness of EFL learners' narrative writing?

d. Does increasing the complexity of a task affect referential cohesion of EFL learners' narrative writing?

e. Does increasing the complexity of a task affect deep cohesion of EFL learners' narrative writing?

2. Does increasing the complexity of a task affect coherence of EFL learners' narrative writing?

Methodology

Research design

A one-group pretest- posttest design, one of the poor experimental designs, in which just a single group is "measured or observed not only after being exposed to a treatment of some sort but also before" was employed in this study (Frankel, Wallen & Hyun, 2012, p.269).

Participants

Forty-one freshmen (33 female and 8 male students) studying at the ELT department of a state university in Turkey and whose ages ranged between 19-28 years participated in the study. The proficiency levels of students were generally intermediate whereas there were also a few students at an advanced level, as also understood particularly from their scores of general writing achievement. Moreover, although 35 students were in their second year as they had prep-class (included writing course) in the previous year, 6 of the participants were newcomers who did not take a writing course before. Before collecting data, basic training for writing essays was provided within the scope of the writing course. Before performing the tasks involved in this study, students were asked to write paragraphs and essays to assess their levels and proficiency in writing. Based on the results, it was regarded that students were homogenous in terms of writing proficiency.

Each participant was involved in narrative writing tasks–simple and complex. In all, 82 essays were involved in the analysis process. All participants were requested to sign a consent form allowing their written production to be used for research purposes. Furthermore, ethics committee approval was also obtained to show that this study complied with ethical standards.

Operationalizations of task complexity

Task complexity was operationalized at two levels as simple and complex based on one of the resource-dispersing variables of Robinson's the Triadic Componential Framework, +/- task structure. Accordingly, the first writing task was identified as complex as a consequence of loose task structure which required more cognitive demand in performance of the task. The second task carried out two weeks later was described as simple since it involved a tight task structure and thus was assumed to need less cognitive demand to design and perform the writing process.

Data collection procedure

Data were collected during a writing course for students studying at the ELT department of a state university. Before starting to collect the data for the study, the instructor, also one of the researchers of this study, provided them with basic education about the writing process and asked them to write some sample paragraphs based on the knowledge they gained and one essay considering the features of an effective essay. After learning the features of narrative writing theoretically, the students were required to produce a narrative essay. Accordingly, they were first shown a picture and asked to analyze it for five minutes. They were then given an hour to create a story based on the picture and thus perform their complex writing task (-task structure/-TS). In another lesson, the students were given a sheet involving 16 related pictures in order and create a story following the order of the pictures for the simple task (+ task structure/ +TS).

Data analysis

Text easibility

The texts produced by the students were analyzed by the Coh-Metrix Common Core Text Ease and Readability Assessor (T.E.R.A.) which is designed to evaluate the easibility and readibility of texts and thus provide valuable information about the characteristics of effective texts. Five components of text easability used in this study were briefly explained below (McNamara et al., 2014, p. 85):

Narrativity: It deals with whether the text is story-like and involves familiar words, world knowledge, and conversation. Therefore, the higher a text is in narrativity, the easier it is.

Syntactic simplicity: It "reflects the degree to which the sentences in the text contain fewer words and use simpler, familiar structures that are less challenging to process". It is measured based on a variety of indices such as the number of clauses for sentences, the number of words in each sentence, and the number of words used before the main verb of the clause. Accordingly, the texts with fewer words or clauses in sentences are syntactically simple and thus easy to comprehend.

Word concreteness: It shows whether the words are concrete and meaningful which helps the reader to visualize and comprehend the text.

Referential cohesion: Referential cohesion in a text illustrates the extent of connections that link the ideas together to help the reader process easily.

Deep cohesion: A text with high deep cohesion contains causal and intentional cohesive devices where they are required to establish relationships among ideas, events, or actions since these devices enable the reader to understand the relationships more deeply and coherently.

Coherence

Since coherence is a subjective feature of writing and simply what the reader grabs from the text (Crossley et al., 2016), we analyzed it through the same analytic rubric also used for the analysis of the students' general writing achievement. The rubric we used included a separate section to evaluate coherence under the name of organization and coherence. As in other sections of the rubric, the scores ranged between 1 and 5. If a text takes the maximum score of 5 for coherence, it means that the text uses a logical structure regarding the purpose, audience, subject of the paper, utilizes true and enough transitions to build a clear connection between sentences, and lead the reader to comprehend the chain of reasoning or progression of ideas. On the contrary, if it is given the minimum score of 1, the text lacks organization, coherence, and transitions.

Statistical Analysis

In order to compare the results of the simple task and complex task performed by the same learners, data were firstly analyzed in terms of normality. Since they were not normally distributed, a Wilcoxon signed-rank test, a nonparametric test, was applied to assess the difference between the two tasks in terms of dependent variables, text easibility indices and coherence. When a statistically significant difference was obtained, its effect size was also calculated (Cohen, 1988)

Reliability and Validity

In the light of the literature, it was agreed to use T.E.R.A that provides an automated evaluation of a text. Furthermore, the validation of indices in this program was verified by McNamara et al. (2010) that those indices measure what is expected to be measured and can be compatible with all types of data regarding human performance.

The other dimensions in the current study, coherence not rated by the computer analysis was also evaluated by another rater to ensure the reliability of coding data. Firstly, three raters were trained about the analytical rubric used in the evaluation and then asked to evaluate 30 essays chosen at random to see whether there was consistency between their results. Reliability test results showed a high level of reliability (Cronbach's alpha coefficient=.87). Based on the results of the reliability test showing consistency in three raters' results, two of the raters completed rating the rest of the essays. The overall results of the two raters were again tested for reliability. Although a high level of inter-rater reliability (Cronbach's alpha coefficient=.88) was found, the essays for which raters had two or more-point-difference in rating scores were determined and discussed by the researchers. As a result of the reevaluation of these essays by the two raters, high-level inter-rater reliability (.89) with a Cronbach's alpha coefficient was reached.

Results

The results of a Wilcoxon signed-rank test conducted to see whether increasing the complexity of writing tasks along the task structure led to a significant difference in the text easibility indices and coherence of EFL learners' narrative writing is displayed in tables and explained in detail under each dependent variable. The analysis results in T.E.R.A. that provide some explanation for these indices explained in the previous section under the subheading of data analysis are also benefited to interpret the results in the tables.

Narrativity

Table 1. Wilxocon signed-test results for the narrativity of the tests

	Narrativity for the simple task - Narrativity for the complex task
Z	-.57
Asymp. Sig. (2-tailed)	.57

According to the results illustrated in Table 1, the difference between complex and simple tasks is not statistically significant. It is also clear that the narrativity of students' narrative essays is not affected by the increase in the cognitive complexity of the tasks (Z=-.57, p=.57>0.05). Indeed, the median narrativity rating is higher in the complex task (median=95.0) compared to that in the simple task (median= 96.0).

Accordingly, it is clear that texts high in narrativity are more story-like and probably have more familiar words. Therefore, it is typically easier to understand those texts. In line with the statistical results, it can be concluded that the narrative essays produced based on a picture in a loose task structure (complex task) are more-story like and thus easier to understand

Syntactic Simplicity

Table 2. Wilxocon signed-test results for the syntactic simplicity

	Syntactic Simplicity for the simple task - Syntactic Simplicity for the complex task
Z	-1.77
Asymp. Sig. (2-tailed)	.08

Table 2 illustrates the results for the difference between two tasks in syntactic complexity of essays produced by the learners. Although the essays produced in the simple task (median=88) had more syntactic

simplicity than those in the complex task (median=83) did, the difference was not found statistically significant (Z=-1.77, p=.08>0.05).

Accordingly, students produced more syntactically simple essays in their simple tasks. In other words, narrative essays in the simple task had more simple sentence structures and thus are easier to process.

Word Concreteness

Table 3. Wilxocon signed-test results for the word concreteness

	Word Concreteness for the simple task - Word Concreteness for the complex task
Z	-3.32
Asymp. Sig. (2-tailed)	.00

A Wilcoxon signed-rank test the result of which is displayed in Table 3 shows that the complexity of the task yielded to a statistically significant difference in word concreteness of learners' narrative essays (Z=-3.32, p=.00<0.05, r=0.51). According to the results of this test, students used more concrete words in their essays they produced in the simple task (median=83) than they did in their essays produced in the complex tasks (median=61). The complexity of the task affects word concreteness at a significance level. It can be also concluded that the use of more concrete words in a simple task makes their narrative essays easier to visualize and comprehend.

Referential Cohesion

Table 4. Wilxocon signed-test results for the referential cohesion

	Referential Cohesion for the simple task - Referential Cohesion for the complex task
Z	-2.22
Asymp. Sig. (2-tailed)	.03

According to Table 4, referential cohesion in the written production of the simple task significantly differs from that in essays of the complex task (Z=-2.26, p=.03<0.05, r=0.34). In addition, the simple task essays had stronger referential cohesion (median=60) compared to the essays of the complex task (median=46). The effect of the task complexity on referential cohesion of written production is statistically significant at a medium level.

Based on these results, it is clear that students' narrative essays in the complex task had little overlap in words and ideas and thus required the reader to make inferences. Therefore, these essays are more difficult to comprehend than those produced in a simple task.

Deep Cohesion

Table 5. Wilxocon signed-test results for the deep cohesion

	Deep Cohesion for the simple task - Deep Cohesion for the complex task
Z	-1.22
Asymp. Sig. (2-tailed)	.22

As shown in Table 5, the difference between the complex and simple tasks in deep cohesion was not statistically significant (Z=-1.22, p=.22>0.05). Although students produced narrative essays in their complex tasks which are richer in deep cohesion (median=78) compared to their essays in the simple task (median=70). In this sense, essays in the complex tasks can be said to have more connecting words to clarify the relationships between events, ideas, and information and thus are easier to comprehend due to this added support.

Coherence

Table 5. Wilxocon signed-test results for the coherence

	Coherence for the simple task - Coherence for the complex task
Z	-.83
Asymp. Sig. (2-tailed)	.41

Although the components of text easibility were analyzed by an automated program, T.E.R.A, essays were assessed by two different raters in terms of coherence using an analytic rubric. Since there was interrater reliability between the scores of the raters, the mean of their scores were taken and statistically analyzed. According to the results displayed in Table 6, there is no statistically significant difference between narrative essays of students in the simple task and those in the complex task in terms of coherence (Z=-2.36, p=.41>0.05). Indeed, the median coherence rating was 3.5 for both complex and simple tasks. Therefore, the complexity of the tasks does not affect the coherence of the texts produced by the learners.

Discussion and Conclusion

This study was investigated to see whether increasing the complexity of a writing task had a significant effect on the easibility of the texts learners produced as a result of their task performance and also on its coherence. Two writing tasks were designed at two levels as complex and simple according to +/- (tight/loose) task structure, one of the resource-dispersing variables in Robinson's framework. Therefore, the participants of this study were asked to produce a narrative essay for the complex task involving one picture about which they were required to narrate a story and also another one for the simple task involving 16 pictures on which they would ground their narration according to the order of the pictures.

Their essays were analyzed by an automated program, T.E.R.A, for the text easibility including five indices as narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion, and by an analytic rubric for the coherence of the texts. T.E.R.A. suggests that the texts higher in these indices are easy to comprehend.

The results of this study pointed out that students' narrative essays in the simple task had more syntactic simplicity although they were low in narrativity. That is, their essays in the complex task were more story-like and thus easier to comprehend but had more complex structures than their essays in the simple task did. However, although differences in these indices were not statistically significant, the complex task had more syntactically complex texts. Similarly, the studies, based on the Trade-off Hypothesis, found that increasing the complexity of the task led to the appearance of more syntactically complex structures in task performance (Salimi et al., 2011; Tavakoli, 2009). However, Tavakoli and Foster (2008) pointed out that syntactic complexity was high in narrative tasks with background information which was identified as the simple task. Furthermore, some studies obtained similar results that task complexity did have no significant effect on the syntactic complexity of the task outcome (Adams et al., 2015; Frear & Bitchener, 2015).

The current study also found that the texts in the simple task had more concrete words, which enabled the reader to visualize what was said and thus to easily comprehend them. However, their narrative essays produced in the complex task involved more abstract words which are assumed to be more difficult to comprehend. These results seemed to support the results of Frear and Bitchener (2015) showing that the use of complex words was positively related to the complexity of the task. Furthermore, Ong and Zang (2010) revealed that increasing the complexity of the task along the provision of ideas and macro-structure affected lexical complexity of EFL learners' argumentative writing whereas task complexity along draft availability had no impact on lexical complexity. Similarly, some studies provided contradictory results that the lexical complexity of participants' task performance was not influenced by the task complexity (Kuiken & Vedder, 2007).

In addition, students' narrative essays in the simple task had more overlap in words and ideas which helped the readers to make inferences and thus easily understand the text. On the other hand, the essays produced in the complex task were richer in deep cohesion and had more connecting words to clarify the relationships between events, ideas, and information although the difference between the two tasks in deep cohesion was not significant. Similarly, no significant difference was observed in the coherence of the texts produced in the simple and complex tasks. It is not so easy to discuss these results in the light of other studies since no study investigating task complexity in terms of cohesion and coherence was encountered. However, when coherence and cohesion are regarded as components of fluency in writing, the studies investigating task complexity in terms of +/- planning suggested that planning time in task positively affected the fluency of written production (Ellis & Yuan, 2004; Ong & Zang, 2010).

Based on these results, clear conclusions cannot be suggested about the effect of task complexity on text easibility since the results show differences in indices. Whereas the essays in the simple task were higher in some indices suggesting that they were easier to comprehend, the essays produced in the complex task had higher scores in other indices. However, it was seen that students produced easier texts in simple tasks, though not statistically significant in all indices. Therefore, conducting more writing tasks at different levels of task complexity may yield more reliable and concrete results for the text easibility. Furthermore, other measures such as complexity, accuracy, and fluency concurrently in addition to those involved in this study can be employed to have further insight into the effects of task complexity on written task performance..

Acknowledgments or Notes

References

Adams, R., Nik Mohd Alwi, N. A., & Newton, J. (2015). Task complexity effects on the complexity and accuracy of writing via text chat. *Journal of Second Language Writing, 29*, 64-81. https://doi.org/10.1016/j.jslw.2015.06.002

Ahmadian, M. J., Abdolrezapour, P., & Ketabi, S. (2012). Task difficulty and self-repair behaviour in second language oral production. *International Journal of Applied Linguistics, 22*(3), 310-330. https://doi.org/10.1111/j.1473-4192.2012.00313.x

Arslanyilmaz, A. (2013). Computer-assisted foreign language instruction: task based vs. form focused. *Journal of Computer Assisted Learning, 29*(4), 303-318. https://doi.org/10.1111/jcal.12003

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing, 32*, 1-16. https://doi.org/ 10.1016/j.jslw.2016.01.003

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. https://doi.org/10.1093/applin/amp042

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. https://doi.org/10.1017/S0272263104261034

Frankel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. Mc Graw Hill.

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

Genç, Z. S. (2012). Effects of strategic planning on the accuracy of oral and written tasks in the performance of Turkish EFL learners. In A. Shehadeh & C. A. Coombe (Eds.), *Task-Based Language Teaching in Foreign Language Contexts* (pp. 67-88). John Benjamins Publishing.

Halliday, M. A. K., & Hassan, R. (1976). *Cohesion in English*. Longman.

Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning, 63*(2), 330-367. https://doi.org/10.1111/lang.12008

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Planning and task performance in as second langauge* (pp. 143-164). Amsterdam: John Benjamins.

Kim, N. (2020a). The effects of different task sequences on novice L2 learners' oral performance in the classroom. *Language Teaching Research*, https://doi.org/ 10.1177/1362168820937548

Kim, N. (2020b). Reconsidering task complexity through different planning and proficiency in l2 written tasks. 응용언어학, *36*(1), 65-95. https://doi.org/ 10.17154/kjal.2020.3.36.1.65

Kırkgöz, Y. (2014). Task-based language teaching. In S. Çelik (Ed.), *Approaches and principles in English as a foreign language (EFL) education*. Egiten Kitap.

Knoch, U. (2007). 'Little coherence, considerable strain for reader': A comparison between two rating scales for the assessment of coherence. *Assessing Writing, 12*(2), 108-128. https://doi.org/10.1016/j.asw.2007.07.002

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20*(2), 148-161. https://doi.org/ 10.1016/j.jslw.2011.02.001

Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster-Cohen, M. d. P. Garcia Mayo, & J. Cenoz (Eds.), *Eurosla Yearbook* (Vol. 5, pp. 195-222). John Benjamins.

Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *IRAL - International Review of Applied Linguistics in Language Teaching, 45*(3). https://doi.org/10.1515/iral.2007.012

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*(1), 48-60. https://doi.org/10.1016/j.jslw.2007.08.003

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (Third Edition ed.). Oxford University Press.

Lee, I. (2002). Teaching coherence to ESL students-a classroom inquiry. *Journal of Second Language Writing, 11*, 135-159. https://doi.org/160-3743/02/$

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2009). Linguistic Features of Writing Quality. *Written Communication, 27*(1), 57-86. https://doi.org/10.1177/0741088309351547

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes, 47*(4), 292-330. https://doi.org/10.1080/01638530902959943

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences-A classroom-based study. *The Modern Language Journal, 95*(Supplementary Issue), 162-181. https://doi.org/10.1111/j.1540-4781.2011.01241

Richards, J. C., & Rodgers, T. (2001). *Approaches and methods in language teaching*. Cambridge University Press.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57. https://doi.org/ 10.1093/applin/22.1.27

Robinson, P. (2003). The Cognition Hypothesis, task design and task-based language learning. *Second Langauge Studies, 21*(2), 45-107.

Robinson, P. (2005). Cognitive complexity and task squencing: A review of studies in a Componential Framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*(1), 1-33. https://doi.org/ 10.1515/iral.2005.43.1.1

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects of L2 speech production, interaction, uptake and perceptions of task difficulty. *Iral-International Review of Applied Linguistics in Language Teaching, 45*, 161-176. https://doi.org/ 10.1515/iral.2007.009

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *Iral-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176. https://doi.org/ 10.1515/iral.2007.007

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

Salimi, A., & Dadaspour, S. (2012). Task complexity and SL development: Does task complexity matter? *Procedia - Social and Behavioral Sciences, 46*, 726-735. https://doi.org/ 10.1016/j.sbspro.2012.05.189

Salimi, A., Dadaspour, S., & Asadollahfam, H. (2011). The effect of task complexity on EFL learners' written performance. *Procedia - Social and Behavioral Sciences, 29*, 1390-1399. https://doi.org/10.1016/j.sbspro.2011.11.378

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

Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Eds.) *Processing perspectives on task performance (pp. 211-*260). Johns Benjamin Publishing Company.

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

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge University Press.

Tavakoli, P. (2009). Assessing L2 task performance: Understanding effects of task design. *System, 37*(3), 482-495. https://doi.org/10.1016/j.system.2009.02.013

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*(2), 439-473. https://doi.org/ 10.1111/j.1467-9922.2011.00642.x

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language*. John Benjamins.

Todd, R. W., Thienpermpool, P., & Keyuravong, S. (2004). Measuring the coherence of writing using topic-based analysis. *Assessing Writing, 9*(2), 85-104. https://doi.org/10.1016/j.asw.2004.06.002

Watson Todd, R., Khongput, S., & Darasawang, P. (2007). Coherence, cohesion and comments on students' academic essays. *Assessing Writing, 12*(1), 10-25. https://doi.org/10.1016/j.asw.2007.02.002

Willis, J. (1996). *A framework for task-based learning.* Longman.

Yang, W. (2014). *Mapping the relationships among the cognitive complexity of independent writing tasks, L2 writing quality, and complexity, accuracy and fluency of L2 writing.* (Unpublished Dissertation Thesis). Georgia State University, Retrieved from http://scholarworks.gsu.edu/alesl_diss/29/

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing, 28*, 53-67. https://doi.org/10.1016/j.jslw.2015.02.002

Yıldız, M., & Yeşilyurt, S. (2017). Effects of task planning and rhetorical mode of writing on lexical complexity, syntactic complexity, and overall writing quality of EFL writers' task performance. *Journal of Language and Linguistic Studies, 13*(2), 440-464.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1-27.