In the Name of God

Journal of Language Horizons

Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Publisher: **Alzahra University**
Executive: **Dr. M. Shirazizadeh**
Chief Editor-in-Chief: **Dr. E. Sotoudehnama**

English Editor: **Dr. Sh. Teimournezhad**
Persian Editor: **Dr. N. Monfared**
Managing Director: **N. Jafari**

Page Layout, Printing and Binding**: Mehrravash Pub**. (+98 91 99 19 92 42)

Alzahra University Publication
Address: Alzahra University, Vanak, Tehran, Iran. Postal Code: 1993891176
Email: lghorizons@alzahra.ac.ir
ISSN**: 2588-350X** E-ISSN**: 2588-5634**

The Journal of Language Horizons is published by **the Iranian Linguistic** Association

the Iranian Linguistic Association

Journal of Language Horizons

According to the Ministry of Science, Research and Technology proclamation dated 1397/12/06 and numbered /3/18/311330, the *Journal of Language Horizons* was granted full Academic-Research status by the ministry's publication committee in their meeting of Jan. 30, 2019. This status has been applied from Volume 1, Issue 1, Spring/Summer 2017.

It is a monolingual (English) journal which operates on a blind peer review policy; it publishes twice- yearly (biannually) papers which report the findings of original research on the current trends and topics in different language related issues.

Submission of manuscripts should be sent electronically via Journal website:
http://lghor.alzahra.ac.ir.

More information about guidelines for submission and the policy of the Journal is provided in other parts of the website. Only articles received online through the site will be sent to referees.

Office Address: Journal of Language Horizons, 2nd Floor, Faculty of Literature and Languages, Alzahra University, Tehran, Iran

E-mail: lghorizons@alzahra.ac.ir

Aims and Scope

Considering the National Master Plan for Science and Education in the country and the necessity of *localizing* the humanities, this journal welcomes English articles contributing to this aim.
Aims of this journal include but are not limited to:

- Promoting scientific and professional knowledge of researchers in the below mentioned areas.
- Sharing the findings of researchers in the mentioned fields.
- Promoting general and professional knowledge necessary for learning English as an international language among Iranians and international interaction in the future.
- Promoting Iranian culture along with the English.

The main focus of the Journal is on research conducted on language learning and teaching. We are in particular interested in research papers on L2 education (in particular EFL/ESL), in a variety of levels, including school level, university level, institute level, etc. The Journal welcomes papers on teaching and learning any component of language including skills and sub-skills, as well as sociolinguistics, psycholinguistics, discourse and pragmatics, TESL or TEFL, Teaching Persian to Speakers of Other Languages (AZFA), language evaluation/testing, language and culture, teaching language for specific purposes, teaching methodology, English and Persian literature, language studies, translation, and linguistics but written only in English.

Submission Guidelines

Authors are encouraged to submit complete unpublished and original works which have not already been published elsewhere nor are under review in any other journal. They are kindly requested to refer to the Journal website: http://lghor.alzahra.ac.ir/ to register according to the instructions given. Once registered, each author will be provided with a username and a password enabling him/her to follow up the submission process. The next step is to send the paper to the Journal to be considered for review.

The manuscripts submitted to *Journal of Language Horizons* should follow the latest version of APA style. Contributing authors are advised to read this document carefully and adhere to the instructions given below before sending their papers to the Journal office. The article will be sent in two separate files including the title file and the manuscript file. The title file includes the title, author's

/authors' name(s), their institutional affiliation(s), phone number(s), and the email(s), preferably the academic one(s). In case of multiple authors, the corresponding author should be indicated.

The second file, the main manuscript, starts with the abstract. The abstract should be between 150 and 250 words, followed by five keywords, separated by commas. The abstract should include information on the purpose of the research and/or research question(s), the methods and materials used, information on the analysis procedures as well as the major findings. The Persian translation of the abstract should be provided too. The font type should be Times New Roman and the size is 12. The whole manuscript should be double-spaced throughout and the new paragraphs should be indented. The manuscript should be divided into clear sections such as: Introduction, Review of literature (which may include sub-sections), Method (including participants, materials, and procedure), Results, Discussion, Acknowledgements (if necessary) and References (and Appendices, if needed). The reference list should be on a new page, double spaced, and the hanging indent method should be used (all lines after the first one are indented). The length of the paper should be between 7000 and 10000 words. The following pages provide key information and give examples of APA style. More information on APA can also be found in the website: http://www.apastyle.org/

REFERENCES

All in-text citations must be listed in full in the reference list at the end of each article following the specifications of the *APA* manual and all references listed must be cited somewhere in the text. Begin the reference list on a separate page entitled "References" and double-space it throughout. Each entry must include the author's name, co-authors (if any), publication date, and title of work. For a journal article, also provide the name of the journal, volume and issue numbers, and page numbers for the article. For an article in an edited volume, list the editor's name, title of the collection, and page numbers of the article. For a book or monograph, list the edition, volume number, series, and name of publisher. For all online works, the DOI should be provided as well. If not available, the URL can be provided instead. Punctuate and capitalize as in the following examples:

Book

Brown, G., & Yule, G. (1989). *Discourse analysis*. Cambridge University Press.

Journal article

Bacon, S. M., & Finnemann, M. D. (1990). A study of the attitude, motives, and strategies of university foreign language students and their disposition to authentic oral and written input. *The Modern Language Journal*, *74,* 459-473.

Journal article with volume number

Lee, M-B. (2002). A closer look at language learning strategies and EFL performance. *Foreign Languages Education*, *10*(1), 115-132.

Article in a book

Dechert, H. (1983). How a story is done in a second language. In C. Faerch & G. Kasper (Eds.), *Strategies in inter-language communication* (pp. 175-195). Longman.

Journal Article with DOI

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, *3*, 635-647. http://doi.org//10.1037/0278-6133.24.2.225

Journal Article without DOI (when DOI is not available)

Hamfi, A. G. (1981). The funny nature of dogs. *E-journal of Applied Psychology, 2*(2), 38-48. http://www.ojs.lib.swin.edu.au/index.php/fdo

Encyclopedia Articles

Brislin, R. W. (1984). Cross-cultural psychology. In R. J. Corsini (Ed.), *Encyclopedia of psychology* (Vol. 1, pp. 319-327). Wiley.

Developmental genetics. (2005). *In Cambridge encyclopedia of child development*. http://www.credoreference.com.library.muhlenberg.edu:80/entry/cupchilddev/developmental_genetics

Reference citation in text

Direct citation 1
She states, "the 'placebo effect' ... disappeared when behaviors were studied in this manner" (Miele, 1993, p. 276), but she did not clarify which behaviors were studied.

Direct citation 2: Fewer than forty words
Lee (1999) found that "The EAP writing curriculum incorporate reading and analysis of major academic journal articles in the specific field to identify macro-level organization an obligatory 'moves' in conjunction with writing practice." (p. 21)

Direct citation 3: More than forty words
Miele (1993) found the following:
The "placebo effect," which had been verified in previous studies, disappeared when behaviors were studied in this manner. Furthermore, the behaviors were never exhibited again, even when reed rings were administered. Earlier studies were clearly premature in attributing the results to a placebo effect. (p. 276)

One work by two authors
- Smith and Takamoto (1997) argued that. ...
- In recent study of SLA (Smith & Takamoto, 1997) ...

One work by multiple authors: Use et al.
- First citation: Jones et al. (1997) found that...
- in a recent study of second language acquisition (Jones et al, 1997)....
- Subsequent citations: Jones et al. (1997) further argued that. ...
- In a recent study of second language acquisition (Jones et al., 1997)...

One work by multiple authors, more than 6
- First citation: Dresler et al. (1992) showed that...
- In a recent study of second language acquisition (Dresler et al., 1992)...

Multiple works: Arrange in alphabetical order
Previous research (Lass, 1992; Meyer & Sage, 1978, 1980; Nichols, 1987a, 1987b; Oats et al., 1973)...
For works not included in the above examples refer to *Publication Manual of the American Psychological Association* (7th ed.).

Review Policy

Since the Journal uses a double-blind review system, it is essential that all author identifying information be removed from the paper and that author(s) information should only be provided in the title page. In case the author's/authors' work is mentioned in the manuscript, replace the name(s) with *Author(s)* in the the text. In other words, there should be no trace of the author in the text.

Before submitting your article to be considered for publication in the Journal of *Language Horizons*, make sure that it adheres to all guidelines given in this document; otherwise, the Journal office may reject the paper before sending it out for review. The papers following these general criteria may be submitted to be evaluated for possible publication. Please be informed that the Journal takes plagiarism very seriously and the contributors are cautioned against this.

The journal follows a rigorous reviewing policy. Each submitted paper is first evaluated for its style consistency and appropriateness of the topic. If found faulty or not appropriate, the paper is returned to the corresponding author for further work and resubmission. The papers that meet initial submission criteria are then reviewed by members of editorial and advisory board as well as external review-ers. A final decision is made on the status of the paper based on the feedback offered by board members and anonymous reviewers to the Editor in Chief. The final decision will be in form of 1) Accepted as it is; 2) Minor revisions; 3) Major revisions; and 4) Rejected. The Journal keeps the right of literary and technical changes but not changes in the content of the articles. Also, article publishing depends on the editorial board's approval.

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Contents

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

An Investigation into Textual Input Enhancement and Output Production in Relation to the Noticing of English Relative Clauses: The Case of Iranian EFL Learners

Research Article
pp. 7-27

Somayeh Sadeghi[1]
Parviz Maftoon*[2]
Masood Yazdanimoghadam[3]

Received: 2020/06/13 Accepted: 2021/02/06

Abstract
Over recent decades, second language acquisition (SLA) researchers have considered attention an important cognitive process mediating second language (L2) learning. The concept of noticing as conscious attention to form has become especially important in the field. The present study explored whether two pedagogical interventions, namely textual input enhancement (TIE) and learners' output production (LOP), in isolation and in combination, promoted learners' noticing and learning of English relative clauses (RC). The study was conducted with a sample of 113 freshmen majoring in English language and literature. The participants were assigned to three experimental groups and the control group based on the treatments they received, +TIE+LOP, +TIE-LOP, -TIE+LOP, and -TIE-LOP. Before treatment, all the participants took a test of English RCs as the pretest. Then, the participants were required to read the reading texts including instances of RCs and take note. While the participants in the +TIE groups read the enhanced input, the participants in the -TIE groups received the unenhanced input. Also, the participants in the +LOP groups were required to carry out output tasks, but the participants in the -LOP groups answered comprehension questions. Consequently, the participants took part in think-aloud processes. Finally, the test of RCs was administered again as the posttest. The findings revealed that LOP positively affected the participants' noticing and learning of RCs. However, TIE was only effective in promoting the participants' noticing, but not learning, of RCs.

keywords: attention, learning, learners' output production, relative clauses, textual input enhancement

* Corresponding Author
[1] Ph.D. Candidate, Department of Language and Literature, Faculty of Humanities and Social Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran. somsadeghi04@gmail.com
[2] Associate Professor, Department of Language and Literature, Faculty of Humanities and Social Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran. pmaftoon@srbiau.ac.ir
[3] Assistant Professor, English Department, Garmsar Branch, Islamic Azad University, Garmsar, Iran. mym1300@gmail.com

DOI: 10.22051/lghor.2021.32221.1334

Introduction

Attention as a cognitive process mediating learning has received special interest in the field of second language acquisition (SLA) in the past decades. "As the field of SLA research enters the new millennium, it is becoming increasingly clear that the role of attention is as important to understanding SLA as the role of input" (Simard & Wong, 2001, p. 104). Attention is viewed as a crucial concept for establishing how language input is processed and then internalized by L2 learners (Hama, 2012). Several theoretical claims have been posed for the essential role of attention in SLA (e.g., Robinson, 1995; Schmidt, 1990, 2001; Tomlin & Villa, 1994; VanPatten, 2004). While there is a general consensus on the crucial role of attention, disagreement exists regarding the amount and type of attention needed for learning (Izumi, 2002).

Schmidt's (1990, 2001) studies on consciousness, awareness, and attention led to the emergence of the Noticing Hypothesis, claiming that "intake is that part of the input that the learner notices" (Schmidt, 1990, p. 139). As such, Schmidt views noticing, which needs learners' awareness and focal attention, as a necessary condition for L2 learning. However, Schmidt (2001) posits a weaker claim, stating that "people learn about the things that they attend to and do not learn much about the things they do not attend to" (p. 30). In line with Schmidt (2001), considering noticing as the correlate to what psychologists call 'attention', noticing and attention are used synonymously in the present study.

Some pedagogical techniques have been proposed for drawing learners' attention to linguistic items to facilitate L2 learning. Textual input enhancement (TIE) (Sharwood Smith, 1991) is regarded as one of the implicit techniques of focus on form in that the physical appearance of certain elements of a text is typographically modified. The focused form is underlined, bolded, italicized, capitalized, or it is written in a different color, a different font, or a different size. TIE has been the focus of a number of SLA research studies which resulted in mixed findings.

Output production has also been suggested as a way of inducing the learner's noticing of L2 forms by Swain (1985). The Output Hypothesis by Swain (1985, 1995, 2005) postulates that as learners produce L2, they may "notice that they do not know how to say (or write) precisely the meaning they wish to convey" (Swain, 2005, p. 474), that may consequently draw learners' attention to L2 features they need in order to find a solution to their interlanguage problems. However, the noticing role of output has been less explored empirically in SLA research.

Literature Review

In the recent decades, a proliferation of SLA studies have investigated the efficacy of TIE and LOP in triggering learners' noticing of form. The results of the research that has explored the impacts of TIE on learners' noticing and acquisition of linguistic forms vary greatly. A number of studies have revealed evidence for positive impacts of TIE on acquisition (e.g., Doughty, 1991; Jourdenais et al., 1995; Lee, 2007; Simard, 2009), and on noticing (e.g., Alanen, 1995; Izumi, 2002; Winke, 2013).

LaBrozzi (2014) examined the effect of different types of TIE on recognition of L2 form as well as reading comprehension. One hundred and nine English speakers learning Spanish were divided into the experimental groups, each receiving one type of TIE, and a control group. The results showed that TIE

positively affected the target structure recognition, but comprehension was not affected by the enhancement.

Naseri and Khodabandeh (2019) investigated whether audio-visual input enhancement affect acquisition and correct use of collocation. The results of posttests revealed the positive effects of audio-visual input enhancement on the acquisition of collocation and the enhanced accuracy regarding use of collocation in the narrative writing.

However, some studies have revealed no effect of TIE on either noticing or acquisition (e.g., Leow, 1997, 2001; Leow et al., 2003; White, 1998). For example, Loewen and Inceoglu (2016) explored how TIE affected noticing and acquisition of L2 verb forms by college-level Spanish learners. Noticing of the target forms was measured using an eye tracker. The results showed that TIE did not trigger learners' greater noticing of L2 verb forms, nor did it show higher levels of noticing the Spanish verb forms as reported by the learners. In addition, the enhanced group was not significantly different from the non-enhanced group in terms of intake of the L2 forms.

Still other studies have revealed a negative impact for TIE on comprehension (Lee 2007; Overstreet, 1998). Overstreet (1998) explored the impact of TIE as well as content familiarity on learning Spanish linguistic forms. Overstreet failed to find a beneficial effect of TIE or content familiarity on learning the target forms. Yet, a main negative effect was found for TIE on comprehension.

Research conducted on the noticing role of output has generally investigated if output tasks promoted more noticing and learning of linguistic forms compared with non-output tasks. Generally, these studies have resulted in mixed and inconclusive findings.

Izumi (2002) investigated the noticing role of output in comparison with noticing effects of TIE. Noticing was measured through learners' note-taking. The note-score analyses revealed that TIE significantly affected noticing, but output production did not significantly affect noticing. However, the reconstruction-score analyses indicated the output participants' better noticing of the target structure. The test results indicated no significant learning gains in the TIE only group while output production was found to have positive effects on learning the target structure.

Comparing the impacts of output production with the impacts of input enhancement, Russell (2014) investigated the noticing function of output and whether the noticing and acquisition of the Spanish linguistic forms were related. To measure noticing, the participants underlined the words they considered important during the reading activity. The results showed that output production positively affected noticing of the targeted forms and that pushed output followed by input enhancement improved the participants' learning of the targeted forms. Yet, exposing to input enhancement did not result in acquisition of the target forms.

Amini et al. (2019) investigated what L2 learners noticed throughout their own output and exposure to model texts and how this noticing affected their short-term and long-term writing performance. Leaners' noticing of linguistic forms was measured through note-taking during the writing task. The findings showed that noticing improved the learners' grammatical accuracy in that the learners retained the lexical and grammatical features compatible with their own noticing. It was concluded that the learners' output production promoted their noticing of the required linguistic features.

However, Izumi et al. (1999) did not confirm the noticing role of output

production. The participants in the experimental group followed an input-output-input-output treatment in that they read and underlined a short passage containing instances of English past hypothetical conditional. The results showed that the output group was not significantly different from the non-output group in terms of noticing of the targeted form.

The above-mentioned studies have reported mixed results on the impact of TIE and LOP on L2 learning. Such mixed results make it hard to generalize the effectiveness of TIE and LOP. In addition, in the TIE research, the first assumption underpinning TIE, i.e., inducing learners' noticing of the targeted form, "has more or less been eschewed" in that the majority of these studies have only used acquisition measures and considered the effectiveness of TIE equal to its efficacy to generate learning (Han et al., 2008, p. 601), and no measure of noticing has been used. According to the noticing hypothesis, for TIE to contribute to L2 acquisition, learners are required to notice the target form; thus, it is essential to indicate that noticing actually occurs (Ellis, 2008). Also, the results of the few studies that have measured noticing were inconclusive, with some studies finding positive effects (Izumi, 2002; Jourdenais et al., 1995; Winke, 2013) and others finding no effects (Leow, 2001) of TIE on noticing.

Moreover, although a large bulk of research has addressed the function of output production in learning L2 in recent years, the noticing function of output has been of relatively little interest. Meanwhile, the few studies conducted to explore the noticing function of output resulted in mixed findings. Also, different task types, different proficiency levels, and different target forms were investigated in these studies. Thus, more studies are needed to throw more light on the role of output tasks in SLA.

Taking account of these gaps in the literature, the present study, drawing on Izumi's (2002) study, incorporates measures of noticing as well as learning measures in exploring the impact of TIE and LOP on acquiring the target structures in Iran's SLA context. More specifically, the impact of TIE and LOP on learners' noticing and learning of English RCs are examined in isolation and in combination. However, the present study is different from Izumi's, employing a different output task which has less been empirically investigated, i.e. text reconstruction cloze task. Also, an important modification to Izumi's (2002) study is that this study provides a different operationalization of noticing through triangulated data collection. A current challenge to the studies on noticing is the difficulty in measuring noticing. As no single measure could provide an absolute account of learners' noticing (Uggen, 2012), the present study employs a qualitative measurement as well as the quantitative measurement of noticing in order to tap more precisely into what linguistic forms learners notice. With this end in view, the following research questions are investigated in the present study, with RQ 1 addressing noticing and RQ 2 addressing learning:

RQ1: Do pedagogical interventions TIE, LOP, and TIE+LOP have any statistically significant effect on the learners' noticing of the English RCs?

RQ2: Do pedagogical interventions TIE, LOP, and TIE+LOP have any statistically significant effect on the learners' learning of the English RCs?

RQ3: What do learners notice while comparing their production with the L2 input?

RQ4: What do learners notice while reading the L2 text for comprehension?

Method

Participants

The participants were freshmen majoring in the English language and literature at Payam Noor University of Sari and Islamic Azad University of Ghaemshahr in Iran. The participants, aged 18-34 years, included both males and females. Convenience sampling was used in that intact classes were selected and each class was randomly assigned to the control group or one of the treatment groups. The participants in all groups had similar educational programs. The main consideration was to select the learners who indicated emerging knowledge of English RCs but they had not developed full mastery of the target structures. To ensure that the participants qualified to be included in the data analyses, their knowledge of RCs was assessed through a test of English RCs, which also served as the pretest. The participants whose scores were higher than 90% on the pretest were considered as having mastery on the target structures and were eliminated from the final data analyses. Also, those who scored less than 10% in the receptive section of the test and did no attempts at RCs in the production section were considered to lack developmental readiness for learning the target structures. They were also excluded from the data analyses. Moreover, the participants who did not attend all the treatment testing sessions were not included in data analyses. Finally, one hundred and thirteen learners participated in the study.

Instrumentation

Reading Texts. Eight short reading texts, each including instances of RCs were used in this study. The readability index of all texts, according to Smog Index, was sixth or seventh grade, i.e. easy to read or fairly easy to read. To ensure the participants comprehended the texts, they were required to take reading comprehension tests. The participants in all groups answered more than 90% of the comprehension items correctly.

Testing Materials. A language proficiency test and a test of RCs were used in the study. The Cambridge Standard Key English Test (KET) was used to determine the participants' English proficiency level. KET was considered as an appropriate proficiency test for the participants of this study based on their performance on a range of previous classroom tests of their course books. In addition, the students who acquired higher than 50 (N = 24) and lower than 20 (N = 13) out of 60 on the KET proficiency test were excluded from the study.

To evaluate the participants' knowledge of the targeted structures, the researcher developed a test of English RCs, which served as the pretest as well as the posttest in this study. Two more frequently used tasks in the studies on RCs (e.g., Doughty, 1991; Gass & Mackey, 2007; Izumi, 2002) were included in the test: the grammaticality judgement (GJ) task to evaluate the participants' receptive knowledge and the sentence combination (SC) task to evaluate their productive knowledge of RCs. To ensure the content validity of the test, care was taken to include appropriate number and type of RCs in the test, considering the number and type of RCs used in the reading texts. The reading texts included five types of RCs, classified according to the head noun they modify: Subject type, Direct Object type, Indirect Object type, Object of Preposition type, and Possessive type. Accordingly, the test included these five types of RCs. Also, the same number of each type of RCs were included in the test. The SC task included 15 items, each type of RC represented by three items. The GJ task included 20 items, each type of RC

represented by four items. For each type of RC, one GJ item included the correct form of the relevant RC, and the other three included erroneous RCs.

In addition, four experts specializing in EFL and having over five years' teaching experience at university evaluated the test and approved the content of the test in terms of difficulty level of the items, clarity in wording, use of Standard English and clarity of instruction. Consequently, the test was administered to a group of 8 English majors who were representative of the target population, and the items that included unfamiliar words or expressions as well as those that the learners had problem comprehending were either eliminated or revised. The final test was piloted to another group of 19 freshmen majoring in English language and literature. The internal consistency of each section of the test was calculated separately, using Cronbach's alpha. The GJ test enjoyed a reliability of .79, and the SC test enjoyed a reliability of .87.

In each SC item, the participants were given two sentences which they were required to combine by attaching the second sentence to the first. The results of the piloted test indicated that some learners used such words as *and, or,* and *because* to combine the sentences; therefore, a more guided format of SC applied in Doughty (1991) was used:

The boy is leaving the school. The boy broke the window last week.

The boy ---------- last week is leaving the school.

Each GJ item was a statement including an RC. The participants were required to indicate if the statements were grammatical or ungrammatical. Also, they were required to mark the erroneous parts of the ungrammatical sentences and correct them. Four types of possible errors that RCs may contain were included in the test: nonadjacency of head noun and relative pronoun, incorrect retention of pronoun, using incorrect morphology for relative pronoun, and inappropriate omission of relative pronoun (Doughty, 1991).

Measures of Noticing. The present study used online measures of noticing, namely note-taking (Cho, 2010; Izumi, 2002; Izumi et al. 1999; Izumi & Bigelow, 2000; Song, 2007) and think-aloud protocols (Alanen, 1995; Bowles, 2010; Leow, 1997, 2001; Rosa & O'Neill, 1999). An advantage of online measures is that there is no time period between exposing to L2 input and the reporting; thus, the likelihood of memory decay is reduced and the validity of the data is increased (Gass & Mackey, 2007). The participants were required to take notes when they read the text. Note-taking was used on the grounds that it was compatible with the reading activity and it was considered not to interfere with the activity (Izumi, 2002).

Also, the participants' noticing was measured through think-aloud processes in that the participants self-reported orally what they were thinking about while performing the reading task. Think aloud protocols were considered to make use of the participants' short-term memory and to reach to the data that the participants had already paid attention to (Hama, 2012). Nine participants in each group were engaged in thinking aloud their thought processes. Based on the proficiency test scores, three participants were selected from among the low proficient ones, three participants from the mid proficient ones, and three participants from the high proficient ones.

Furthermore, noticing or absence of noticing was determined based on the participants' attempts at the target structures in the reconstruction tasks. That is, any type of RC attempted and the correct use of the targeted RC in the reconstruction task were considered instances of noticing the RCs.

Text Reconstruction Cloze Task. A text reconstruction task was used as the output task in the present study. A version of text reconstruction, i.e. text reconstruction cloze task, was employed. The reason underlying this decision was the findings of the pilot study, in which most of the participants tended to use structures other than the RCs in free reconstruction of the text; accordingly, it was decided to use a more controlled form of text reconstruction, i.e. text reconstruction cloze task (Nassaji & Fotos, 2014), in order to provide an optimal condition for producing RCs. Reconstruction cloze tasks require learners' ability to understand the context in order to supply the missing words in the text while potentially drawing their attention to the target structures.

Procedure

Before the treatment, the L2 proficiency test and the RC test were administered to all the participants. For RC test, first all the participants took the GJ task in 15 minutes. Upon completion of the GJ, the participants carried out the SC task in 12 minutes. The treatment started one week after the pretest. The treatment for each group took place through 8 sessions, over a 3-week period. In each session, first the participants in all groups read the same reading text. The participants in the +LOP groups took notes on a separate sheet any word they considered important or helpful to reconstruct the text. To prevent the possibility of direct copying, time of exposure was also controlled. To induce their noticing of the targeted structures, the researcher told the participants that they could use their notes during text reconstruction and that the accuracy of sentences in the reconstruction was as important as the content. The participants in the -LOP groups took notes on a separate sheet any word that they considered important to comprehend the text.

After reading the text, the participants in the -LOP groups answered comprehension questions, but the participants in the +LOP groups completed the task of text reconstruction, which aimed to examine the impact of output production in drawing learners' attention to RCs by giving the participants the opportunity for producing the target structures. It was expected that the participants in the +LOP groups would notice the RCs more than the participants in the -LOP groups.

In the second stage of the treatment, the participants read the same text again and took notes as in the first phase. However, in the second phase, the +TIE groups read the text containing enhanced RCs, while the -TIE groups read the same text in which the RCs were left unenhanced. Bolding and capitalizing were used as the techniques of enhancement in the present study in that all words in the RCs were bolded, and all the relative pronouns were also capitalized. Consequently, the participants engaged in the postexposure tasks as in the first phase.

After the second postexposure task, nine participants in each group took part in think-aloud processes. The participants in the +LOP groups compared their reconstructed text with the reading text, located the mismatches, and verbalized their thoughts. If they referred to RCs or parts of RCs as mismatches, they were required to explain the mismatches in order to establish what they were exactly attending to. In the -LOP groups, the participants were asked to read the text once more and think-aloud any parts they considered important to understand the text. When they referred to RCs or parts of RCs as important, they were asked why they considered these parts important in order to establish what they were exactly attending to. The verbal protocols were audio-taped for transcription and identifying instances of noticing RCs. One week after receiving the last session of the treatment, all the participants took the posttest.

Design and Analyses

The present study used both quantitative and qualitative approaches. A pretest-posttest design was followed in the quantitative part of the study. It was quasi-experimental and involved three treatment groups and a control group. The groups varied based on output requirement and exposure to TIE: +TIE+LOP group (N = 29), -TIE +LOP group (N = 28), +TIE -LOP group (N = 28), -TIE -LOP group (N = 28).

The qualitative part of the study used focus groups method to further investigate the effectiveness of the two pedagogical interventions, TIE and LOP, in isolation and in combination. More specifically, nine participants in each group were required to think aloud their thought processes in order to explore 'what' of noticing.

Noticing of the target structures was measured by computing the reconstruction scores and the note scores. The reconstruction scores were arrived at by tallying any type of RC attempted and the correct use of the target RC. Following Cho (2010), Izumi (2002), and Song (2007), the note scores of each participant were computed through dividing the total number of words he/she noted by the number of the target structure-related words in his/her notetaking. Then, a percentage score was calculated in order to minimize variation in the amount of note individuals took.

To compute the test scores, the correct answers were given one point and the incorrect answers were given zero. For each item in the SC part, only when the targeted RC was produced and everything related to the formation of RC was correct, the production was considered correct while errors of spelling and tense were not considered if they were not related to RC formation. Also, relative pronoun omission where it was acceptable was considered correct. For the GJ test, the correct response to each item was given one point. The GJ items received no point when they were judged incorrectly, or when the item was judged correctly, but the erroneous part of the statement was not underlined or corrected.

Results

Probing the First Research Question

As noticing was measured through text reconstruction as well as note-taking for the output groups, -TIE+LOP and +TIE+LOP, the results of the effects of the pedagogical interventions on noticing RCs are reported in two sections below.

Results of the Note Scores. To examine the first research question, a non-parametric analysis of covariance (ANCOVA) was used because, as Table 1 displays, the normality assumption was not retained. Non-parametric ANCOVA was run for comparing the four groups', -TIE-LOP, -TIE+LOP, +TIE-LOP and +TIE+LOP, median scores on the posttreatment note taking of RCs while controlling for possible effect of their initial note taking before treatments.

Table 1
Descriptive Statistics; Testing Normality of Pretreatment and Posttreatment Note Taking of RCs by Groups

Group		N	Skewness			Kurtosis		
		Statistic	Value	Std. Error	Ratio	Value	Std. Error	Ratio
-TIE-LOP	Pretreatment	28	1.618	.441	3.67	4.357	.858	5.08
	Posttreatment	28	1.791	.441	4.06	5.131	.858	5.98
-TIE+LOP	Pretreatment	28	1.233	.441	2.80	1.497	.858	1.74
	Posttreatment	28	1.041	.441	2.36	.618	.858	0.72
+TIE-LOP	Pretreatment	28	.649	.441	1.47	.126	.858	0.15
	Posttreatment	28	.666	.441	1.51	1.802	.858	2.10
+TIE+LOP	Pretreatment	29	.799	.434	1.84	-.226	.845	-0.27
	Posttreatment	29	.031	.434	0.07	-.773	.845	-0.91

Table 2 demonstrates the mean ranks and median scores of the four groups on the posttreatment note taking of RCs. The results indicated the highest mean rank for -TIE+LOP (MR = 81.66) on the posttreatment note taking, followed by the +TIE+LOP (MR = 67.86), +TIE-LOP (MR = 57.55) and –TIE-LOP (MR = 20.54).

Table 2
Mean Ranks AND Median Scores; Posttreatment Note Taking of RCs by Groups

	Group	N	Mean Rank	Median
Posttreatment Note Taking	-TIE-LOP	28	20.54	13.50
	-TIE+LOP	28	81.66	55.00
	+TIE-LOP	28	57.55	42.00
	+TIE+LOP	29	67.86	50.00
	Total	113		

The results of non-parametric ANCOVA (F [3, 109] = 40.99, p = .000), Table 3, showed significant differences between the four groups' mean ranks on posttreatment note taking of RCs after controlling for the possible effects of pretreatment note taking. Therefore, the results indicated that the pedagogical interventions TIE, LOP, and TIE+LOP had statistically significant effects on the learners' noticing of RCs.

Table 3
Quade Nonparametric Analysis of Covariance; Posttreatment Note Taking of RCs by Groups with Pretreatment note taking

F	DFH	DFE	P Value
40.994	3	109	.000

Results of the Reconstruction Scores. To investigate the effects of the pedagogical interventions on the participants' noticing of RCs in the +LOP groups, the text reconstruction task was also considered a measure of noticing. Because of the nominal nature of the reconstruction scores, an analysis of chi-square (crosstabs) was used to compare the -TIE+LOP and +TIE+LOP groups' noticing of RCs in the first and second text reconstruction tasks. As Table 4 displays, the -TIE+LOP and +TIE+LOP groups had almost the same percentages of noticing RCs in the first text reconstruction (38 % vs. 37.7%). They also had almost the same percentages of noticing RCs in the second text reconstruction (62 % vs. 62.3%). The standardized

residual values for first and second text reconstructions were all lower than +/- 1.96; indicating that there were not any significant differences between the two groups on the first and second reconstruction scores.

Table 4
Frequencies, Percentages and Std. Residuals; Text Reconstruction by Group

			Reconstruction		Total
			1st	2nd	
Group	-TIE+LOP	Count	629	1028	1657
		% within Group	38.0%	62.0%	100.0%
		Standardized Residual	.1	-.1	
	+TIE+LOP	Count	628	1038	1666
		% within Group	37.7%	62.3%	100.0%
		Standardized Residual	-.1	.1	
Total		Count	1257	2066	3323
		% within Group	37.8%	62.2%	100.0%

Table 5 indicates the findings of analysis of chi-square. The results (χ^2 [1] = .015, p = .903, Cramer's V = .003, p = .875) indicated there were not any significant differences between the two groups' first and second text reconstruction tasks. Thus, the pedagogical intervention TIE did not have any statistically significant effect on the learners' noticing of RCs.

Table 5
Chi-Square Tests; Text Reconstruction by Group

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.025[a]	1	.875		
Continuity Correction[b]	.015	1	.903		
Likelihood Ratio	.025	1	.875		
Fisher's Exact Test				.886	.452
Linear-by-Linear Association	.025	1	.875		
N of Valid Cases	3323				
Cramer's V	.003				.785

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 626.80.
b. Computed only for a 2x2 table

Probing the Second Research Question

To investigate the second research question, an analysis of covariance was used for comparing the four groups' means on the posttest of RCs. The parametric one-way ANCOVA was run because, as Table 6 demonstrates, the assumption of normality, linearity and homogeneity of regression slopes were retained.

Table 6
Testing Normality of Test of RCs (GJ+SC) by Groups with Pretest

Group		N	Skewness			Kurtosis		
		Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
-TIE-LOP	Pretest	28	.419	.441	0.95	-.854	.858	-1.00
	Posttest	28	.547	.441	1.24	-.916	.858	-1.07
-TIE+LOP	Pretest	28	.209	.441	0.47	-.821	.858	-0.96
	Posttest	28	.149	.441	0.34	-1.283	.858	-1.50
+TIE-LOP	Pretest	28	.304	.441	0.69	-.826	.858	-0.96
	Posttest	28	.315	.441	0.71	-.705	.858	-0.82
+TIE+LOP	Pretest	29	.270	.434	0.62	-.519	.845	-0.61
	Posttest	29	.272	.434	0.63	-1.286	.845	-1.52

Table 7 indicates the descriptive statistics for the four groups on the posttest of RCs after controlling for the effect of the pretest. The findings showed the highest mean scores for the +TIE+LOP group (M = 18.88) on the posttest of RCs, followed by the -TIE+LOP (M = 18.06), the -TIE-LOP (M = 14.97) and the +TIE-LOP (M = 14.15) groups.

Table 7
Descriptive Statistics; Posttest of RCs by Groups with Pretest

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
-TIE-LOP	14.974[a]	.444	14.095	15.853
-TIE+LOP	18.060[a]	.444	17.181	18.939
+TIE-LOP	14.154[a]	.444	13.275	15.034
+TIE+LOP	18.887[a]	.436	18.023	19.751

a. Covariates appearing in the model are evaluated at the following values: Pretest = 14.48.

The findings of one-way ANCOVA (F [3, 108] = 27.32, p = .000, partial eta squared = .431), Table 8, indicates significant differences between the four groups' means on the posttest. Therefore, the results indicated that the pedagogical interventions TIE, LOP, and TIE+LOP had statistically significant effect on the learners' learning of RCs.

Table 8
Tests of Between-Subjects Effects; Posttest of RCs by Groups with Pretest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pretest	6837.050	1	6837.050	1241.268	.000	.920
Group	451.504	3	150.501	27.324	.000	.431
Error	594.877	108	5.508			
Total	38651.000	113				

The results of post-hoc comparison tests (Table 9) showed that:

A: There was no significant difference between the -TIE-LOP (MR = 14.97) and the +TIE-LOP (MR = 14.15) groups' means on the posttest of RCs (Mean Difference = .820, p = .194).

B: The -TIE+LOP group (MR = 18.08) significantly outperformed the -TIE-LOP (MR = 14.97) group on the posttest of RCs (Mean Difference = 3.08, p = .000).

C: The -TIE+LOP group (MR = 18.08) significantly outperformed the

+TIE-LOP (MR = 14.15) group on the posttest of lRCs (Mean Difference = 3.90, p = .000).

Table 9
Pairwise Comparisons; Posttest of RCs by Groups with Pretest

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
-TIE-LOP	+TIE-LOP	.820	.627	.194	-.424	2.063
-TIE+LOP	-TIE-LOP	3.086*	.627	.000	1.843	4.329
	+TIE-LOP	3.906*	.628	.000	2.661	5.150
+TIE+LOP	-TIE-LOP	3.913*	.622	.000	2.680	5.146
	-TIE+LOP	.827	.622	.186	-.405	2.060
	+TIE-LOP	4.733*	.623	.000	3.499	5.967

*. The mean difference is significant at the .05 level.

D: The +TIE+LOP group (MR = 18.88) significantly outperformed the -TIE-LOP (MR = 14.97) group on the posttest of RCs (Mean Difference = 3.91, p = .000).

E: There was not any significant difference between the +TIE+LOP (MR = 18.88) and the -TIE+LOP (MR = 18.08) groups' means on the posttest of RCs (Mean Difference = .827, p = .186).

F: The +TIE+LOP group (MR = 18.88) significantly outperformed the +TIE-LOP (MR = 14.15) group on the posttest of RCs (Mean Difference = 4.73, p = .000).

Probing the Third and Fourth Research Question

To answer the third and fourth research questions, all participants' the think-aloud protocols were transcribed. Two raters coded the think-aloud protocols as a) understanding b) noticing, or c) no report. The interrater reliability was 96%. Three themes emerged from the analysis of think aloud protocols under no report of RCs category: 1) noticing of grammatical points other than RCs, 2) noticing of content words, 3) noticing of unfamiliar expressions and unknown words.

The proportions of features noticed by the participants in all groups are shown in Table 10. The largest proportion of noticing has to do with the content of the reading texts, and the lowest proportion is related to understanding the RCs. Considering the third research question, the results indicated that the output-group participants mostly reported noticing of the RCs in the reading texts while comparing their production with the L2 input, with 44% of the noticing in the -TIE+LOP group and 61% of the noticing in the +TIE+LOP group relating to the RCs.

Regarding the fourth research question, the findings revealed the participants in the nonoutput groups largely reported noticing of the content of the reading texts while reading the L2 text for comprehension, with 40% of the noticing in the -TIE-LOP group and 46% of the noticing in the +TIE-LOP group relating to the content of the reading texts.

Table 10
Percentage of Linguistic Features Reported by Groups

Groups	Understanding	Noticing	General Grammar	Content	Unknown words or expressions
-TIE-LOP	.9	.4	.25	.40	.22
+TIE-LOP	.4	.25	.19	.46	.5
+TIE+LOP	.2	.61	.3	.28	.6
-TIE+LOP	.7	.44	.11	.31	.7
Total	.6	.33	.15	.36	.10

A sample of verbal report including instances of noticing RCs by a participant in the +TIE+LOP group while comparing her output task with the reading input to locate mismatches is given in (1):

(1) here I didn't write the part with 'whom' correctly…. I couldn't write these correctly, parts with 'whom'…. I have problem with who, whom and these, I get confused.

Also, a sample of verbal report from the -TIE+LOP group and a sample of verbal report from the -TIE-LOP including instances of understanding RCs is given in (2) and (3), respectively:

(2) In 'shopkeepers who cannot afford', 'who' is used for 'shopkeepers', we use relative pronoun 'who' for human beings, I wrote it correctly.

(3) Here 'people whose', 'whose' is used after people because 'whose' shows possession.

The analysis of verbal reports that included instances of noticing the grammar issues other than RCs revealed that most of the participants focused primarily on such structures as passive verbs, infinitive phrases, prepositions, gerund phrases, verb phrases including modal verbs, and third-singular 's'. Samples of noticing each of these linguistic features are illustrated in (4-9), respectively:

(4) here, the structure is passive, 'blood pressure is recorded' is passive, the verb is past participle….

(5) 'encourage to buy things' is important. After 'encourage', infinitive is used, 'encourage' is important here….

(6) 'disagree about' is important, its preposition is 'about', after 'disagree' we should use 'about' ….

(7) here, there is gerund, 'of both spending', 'spending' is gerund, I think it is because of 'of', after preposition 'of' gerund was used….

(8) after 'can' the verb is simple, 'can' is modal [verb]. Again here, we have 'may', it is modal [verb], and we have simple verb after it….

(9) this verb has 's' because it is singular, 'social networking' is singular, so the verb comes with 's'….

The verbal reports were classified as 'noticing the content' when the participants referred to the words or expressions that bore the meaning of the input text. Samples of noticing the content are given in (10-12):

(10) it is important that the people lose their control in spending money…that we can see the videos online….

(11) the key points in the text are repeated, they are important…. These sentences have key role in the meaning [of the text] ….

(12) that he could go to the beach…he couldn't do any fun, it is important…. That he was alone, it is important….

Finally, samples of the verbal reports illustrating noticing unknown words are given in (13-14):

(13) 'crawl' and 'resign' I didn't know their meaning, they are important.

(14) here, 'afford' I'm not sure about the meaning, it means pay or something like that, I don't know the meaning.

Discussion

The Effects of the Pedagogical Interventions on Noticing RCs

The first question explored the effects of three attention-drawing pedagogical interventions, TIE, LOP, TIE+LOP, on noticing the English RCs. The findings indicated significant differences among the experimental groups and the control group on noticing RCs after treatments; thus, the efficacy of all pedagogical interventions in promoting noticing of RCs.

The Effect of LOP on Noticing. The positive impact of output indicated in this study is line with the studies that supported the noticing function of output (e.g., Qi & Lapkin, 2001; Izumi, 2002). However, in Izumi's (2002) study, the positive effect of output was found only in the text reconstruction but not in the note-score analyses. Yet, in the present study, both the note-score analyses and the reconstruction-score analyses revealed the beneficial effect of LOP on noticing. Nevertheless, the results of this study are different from studies which did not confirm the effect of output in inducing noticing of the targeted forms (e.g., Izumi & Bigelow, 2000).

The positive impact of output on noticing in this study might be accounted by the type of output task employed in this study. The text reconstruction cloze task used in this study might have been more successful in inducing learners' noticing of the targeted structures than the output tasks used in the studies that failed to find such an effect. Another point that can be accounted for the efficacy of output in the present study regards employing input-output-input technique in that after carrying out the output task, participants were provided with the input text including instances of the target structures. Receiving L2 input *immediately* after output production might further push the participants to focus their attention to the problematic parts in the input, thus maximizing the possibility of noticing the forms in the input (Swain, 1995).

The Effect of TIE on Noticing. The findings also demonstrated a beneficial effect for TIE on noticing the target structures, which are in line with the studies revealing such an effect (e.g., Alanen, 1995; Izumi, 2002; Winke, 2013). The results lend further empirical support to the theoretical rational underpinning input enhancement proposed by Sharwood Smith (1991), positing that enhancing the linguistic features of input, as a way of improving the quality of input, increases the perceptual saliency of the targeted forms and in this way contributes to the increased noticing of the forms.

However, the findings contradict the studies that failed to find a favorable effect of TIE on noticing (e.g., Leow, 1997, 2001; White, 1998). Some factors can be accounted for the positive effect of TIE found in this study in contrast to the studies that were not successful to find such an effect. The first factor regards learner readiness. The participants in this study might have been cognitively and linguistically ready to recognize the target structures enhanced in the reading texts. As Philp (2003) argues, learners might not notice the input which is beyond their current level of L2 development. Philp also refers to learner readiness as the

learner's prior knowledge of and familiarity with linguistic items. The target structures in the present study had been previously introduced to the participants, which is in contrast to previous studies that found no effect of TIE, in which the targeted forms were almost new to the learners (Leow et al., 2003).

Second, the structures targeted in the study carry considerable communicative value. All forms are not equally amenable to TIE techniques; rather, the efficacy of attention-drawing activities to some extent depends on the structure involved (Han et al., 2008). This is evidenced by the studies which indicated that TIE was not effective to draw leaners' attention to the targeted forms with low communicative value (Alanen, 1995; Leow et al., 2003; White, 1998), but it was effective in inducing learners' noticing of the targeted forms carrying substantive communicative value (e.g. Izumi 2002; Winke, 2013).

The third factor that may account the positive effect of TIE regards the technique of TIE used in this study, i.e. bolding and capitalizing. Capitalizing was found to be the most effective TIE technique by Simard (2009). Also, the participants might have been accustomed to bolding technique. As argued by Sarkhosh et al. (2013), Iranian learners typically use bolding as a learning strategy – they usually highlight important parts of their textbooks in order to emphasize those points and to pay more attention to the points in subsequent exposures. The bolding technique resembles highlighting more than other TIE techniques.

Finally, the length of exposure to the target structures can be accounted for the positive effect that TIE had on the participations' noticing. While in a number of former studies, learners were exposed to the enhanced forms for a relatively short time, a single exposure in some studies, in this study, the participants were exposed to the RCs for eight sessions.

Comparing the Effects of TIE and LOP on Noticing. The results indicated that LOP was more effective than TIE in inducing noticing. This is evidenced by the -TIE+LOP group having the highest rank of noticing RCs, as well as the +LOP groups' outperformance on noticing RCs compared with the +TIE only group. Thus, +LOP was found to facilitate greater noticing of the target structures, which is in line with Song (2007) but contradicts Izumi's (2002) findings indicating TIE as a more effective attention-drawing technique than output production.

The first explanation regards the noticing role of output in that the +LOP groups might have experienced some difficulties producing the target structures while performing the first output tasks and they tried to focus more on the target structures in the second exposures to the input texts. However, the +TIE only group did not experience such a heightened need to attend to the target structures after carrying out the task which focused solely on meaning, as Izumi (2002) also acknowledges. This finding suggests that input-output-input technique is more successful than input-only technique in inducing learners' noticing of the target forms.

Another explanation could relate to the difference in the follow-up tasks the participants in each treatment group were required to carry out, which Song (2007) refers to as well. While the +LOP groups were required to reconstruct the texts after the reading activity—input-output-input-output—the +TIE only group read the texts for the sole purpose of comprehension. Since the +LOP groups were required to produce the text, they might have felt greater need to attend to the formal aspects in addition to the meaning as they were told that accuracy in the output tasks was as important as the content. However, for the +TIE only group, who were only required

to carry out the comprehension task, the need to pay attention to the form might have less brought into focus.

The Effects of the Pedagogical Interventions on Learning RCs

The results indicated positive effects of LOP on learning RCs. The +LOP groups indicated a greater intake of RCs compared with the control group. The findings support the results of the studies which found such an effect (e.g., Qi & Lapkin, 2001; Swain & Lapkin ,1995) but in contrast to the studies that failed to indicate positive effect of LOP on learning target forms (e.g., Izumi & Bigelow, 2000).

However, the results did not indicate a beneficial effect of TIE on learning. +TIE only group did not significantly outperform the control group on the posttest of RCs, nor did the +TIE+LOP group significantly outperform the -TIE+LOP group on the posttest. The findings support the studies that failed to find a beneficial effect of TIE on learning (e.g., Leow, 1997, 2001) but contradict the studies that found a positive effect (e.g., Lee, 2007; Simard, 2009).

The findings were different for the two pedagogical interventions regarding intake of RCs. The findings are similar to those of Song (2007), in which the output groups indicated significantly greater intake of the targeted form than the nonoutput groups. Also, the findings are in line with Izumi (2002) who found that output production significantly facilitated learning of the target structures while TIE merely triggered noticing of the targeted structures but did not result in the subsequent cognitive processing needed for learning the structures.

The results of the study regarding the +LOP groups supported the Noticing Hypothesis positing that increased noticing of the form results in increased chance of learning the form (Schmidt, 2001). The +LOP groups, which indicated increased noticing of RCs after treatment, also improved on their performance on the posttest of RCs. Yet, the results on TIE contradicts what was expected based on the Noticing Hypothesis. The +TIE only group, which demonstrated significant increase on noticing RCs after treatment, did not significantly improve on their performance on the posttest of RCs.

One explanation might be the quality of noticing evidenced by the +TIE only group as referred to by Izumi (2002) in that the participants might have not noticed the structures at the level that would lead to deeper processing of the form needed for learning. The quality of attention triggered by LOP and TIE might be different probably because the former induces attention internally, while the latter triggers attention externally; thus, these two interventions "do not promote learning with an equal level of efficacy" (Izumi, 2002, p. 567). The results of the study lend support to the claim that output production engage learners in a deeper processing of form needed for learning than TIE (Izumi, 2002; Song, 2007).

The findings of the study should be interpreted considering a number of limitations. Interpreting the results of this study is restricted by some factors, including contextual factors and individual differences, that might potentially mediate the efficacy of TIE and LOP in promoting noticing of the target structures.

The Qualitative Analysis

The think-aloud protocols' analyses confirmed the findings of the quantitative data in this study, considering the positive impacts of LOP on the participants' noticing of the targeted structures. The +LOP-group participants

verbalized more instances of noticing the RCs than other linguistic forms. The results contradict those of Uggen (2012) in that the stimulated recall protocols in Uggen's study revealed that the vast majority of the noticing episodes did not include noticing of the target structures but other linguistic forms.

However, the beneficial effect of LOP found in the qualitative data analysis in this study is in line with Hanaoka (2007) and Qi and Lapkin (2001), in which the analyses of the think-aloud protocols supported the favorable impact of output on the participants' noticing of linguistic forms in un-planned writing tasks.

Also, the findings of the think-aloud protocols supported the findings of the quantitative data regarding the positive effect of TIE on the participants' noticing of the targeted structures. Although the participants in the +TIE-LOP group reported greater noticing of the content of the reading texts than noticing of the RCs, both +TIE groups were found to notice the RCs more than the -TIE-LOP group did. The results support Alanen (1995) and Jourdenais et al. (1995) in which think-aloud protocols' analyses indicated that TIE promoted noticing of the target forms. Yet, the results differ from Leow's (2001) showing that the enhanced input did not significantly promote noticing of the targeted forms compared to the unenhanced input.

In line with the studies which investigated the nature of noticing linguistic features by L2 learners (e.g., Amini, et al. 2019; Garcia Mayo & Labandibar, 2017; Hanaoka, 2007; Hanaoka & Izumi, 2012), the vast majority of the verbal reports in the -LOP groups in this study involved noticing of the content words and unfamiliar words or phrases.

In addition, the analysis of instances of noticing by the participants revealed that noticing of grammatical points other than RCs was more evident in the -LOP groups, especially in the -TIE-LOP group, probably because in other groups, the treatments, particularly LOP, directed learners' focused attention toward RCs and away from other linguistic forms.

The low proportion of reports on unknown words and unfamiliar expressions in this study can be accounted for by the readability level of the reading texts, which were easy and fairly easy to read; accordingly, they included more familiar words and expressions. Yet, the -TIE-LOP group's greater instances of noticing unfamiliar words can be related to their less focus on linguistic forms in an exclusively meaning-based program, compared to the other groups which were engaged in focus-on-form.

Conclusion

The focus of the study was exploring the impact of TIE and LOP on promoting learners' noticing and learning of English RCs. The findings of both quantitative and qualitative analyses provided evidence for the efficacy of TIE and LOP, in isolation and in combination, in inducing the participants' noticing of the target structures as well as the usefulness of LOP in promoting the participants' intake of the target structures.

As attention is increasingly viewed as the crucial cognitive process mediating input and intake in SLA, triggering learners' attention to linguistic forms can be helpful in developing their interlanguage. Noticing-inducing interventions have been empirically demonstrated to be effective in facilitating L2 learners' intake of linguistic forms in previous research. The particular contribution of the study to the existing body of literature is that it provides practitioners with more insight into

how to employ these pedagogical interventions.

This study incorporated measures of noticing in exploring the impact of TIE, which was disregarded in most of the previous TIE studies. As TIE was shown to be effective only in promoting noticing, but not acquisition, of the targeted structures in this study, it can be concluded that the failure of a number of previous studies to find a beneficial effect of TIE on learning L2 forms might not be due to inefficacy of TIE in inducing learners' noticing, but to inefficacy of TIE to help learners go beyond simple detection of forms and learn the target structures. In other words, TIE might not be a sufficient pedagogical intervention for promoting intake of a number of L2 forms, including RCs, as well as noticing thereof. Thus, teachers are recommended to employ TIE in combination with such other pedagogical interventions as explicit rule presentation or corrective feedback in order to facilitate a level of noticing which would lead to further processing needed for learning.

Another pedagogical implication regards the output task used in this study. The text reconstruction cloze task was shown to be successful in promoting both noticing and learning of the target structures in the present study. Accordingly, teachers and material developers are recommended to employ this rather new output task in L2 classrooms. While in free output production tasks, learners tend to focus dominantly on lexical features (Garcia Mayo & Labandibar, 2017), the findings of the present study, especially the think-aloud protocols, demonstrated that the text reconstruction cloze task, though not purely creative, was successful in providing an optimal condition for triggering the participants' noticing of RCs. Therefore, this version of dictogloss task is suggested to be used as an effective output task through which learners are more likely to use their linguistic resources of the forms in focus, thereby notice their possible deficiencies and develop their interlanguage.

Also, the findings of the think-aloud protocols provide practitioners with some clues on the nature of the linguistic forms more frequently noticed by the L2 learners whom the participants of this study are a sample of. Building on the findings of this study and the similar studies, teachers can design tasks that focus on the grammatical points learners themselves perceive as salient. In this way, they can more efficiently help learners develop their interlanguage, considering that teachers' planned pedagogical focus may not always match what learners actually attend to (Hanaoka, 2007). In other words, the nature of learner-initiated focus on form can guide teacher-initiated focus on form.

Finally, the positive impact of TIE+ LOP on noticing and learning of RCs found in this study adds one piece of evidence to the effectiveness of TIE in combination with LOP and provide practitioners with a variety of pedagogical interventions which could promote L2 learning.

A less investigated issue in the field of research on noticing is to what extent individual differences and contextual factors mediate attentional resources and the impact of noticing-inducing interventions. Further studies are invited to explore learners' motivation, learning style, pattern recognition ability, memory span as well as such structural features as the complexity and the communicative value of the targeted forms in relation to noticing.

Furthermore, future research could shed more light on the efficacy of the text reconstruction cloze task, targeting other linguistic forms and/ or comparing text reconstruction cloze task with other output tasks.

References

Alanen, R. (1995). Input enhancement and rule representation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259-302). University of Hawaii Press.

Amini, D., Amini, M., & Naseri Maleki, F. (2019). Investigating noticing in narrative writing tasks and its effect on EFL learners' writing performance. *Applied Research on English Language, 8*(3), 365-382. http://doi.org/10.22108/are.2019.115769.1437

Bowles, M. (2010). *The think-aloud controversy in second language research.* Routledge.

Cho, M. Y. (2010). The effects of input enhancement and written recall on noticing and acquisition. *Innovation in Language Learning and Teaching, 4*(1), 71-87. http://doi.org/10.1080/17501220903388900

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of relativization. *Studies in Second Language Acquisition, 13*(2), 31-469. https://doi.org/10.1017/S0272263100010287

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

Garcia Mayo, M. D. P., & Labandibar, U. L. (2017). The use of models as written corrective feedback in English as foreign language (EFL) writing. *Annual Review of Applied Linguistics, 37*, 110-127. https://doi.org/10.1017/S0267190517000071

Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research.* Erlbaum.

Hama, M. (2012). *Strategic planning, recasts, noticing, and L2 development* [Unpublished doctoral dissertation]. Georgetown University.

Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input. *Applied Linguistics, 29*(4), 597-618. https://doi.org/10.1093/applin/amn010

Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research, 11*(4), 459-479. https://doi.org/ 10.1177/1362168807080963

Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing, 21*(4), 332-47. https://doi.org/10.1016/j.jslw.2012.09.008

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL Revitalization. *Studies in Second Language Acquisition, 24*(4), 541-577. https://doi.org/10.1017/S0272263102004023

Izumi, S., & Bigelow, M. (2000). Does output promote noticing in second language acquisition? *TESOL Quarterly, 34*(2), 239-278. https://doi.org/10.2307/3587952

Izumi, S., Bigelow, M., Fujiwara, F., & Fearnow, S. (1999). Testing the output hypothesis. *Studies in Second Language Acquisition, 21*(3), 421-452. https://doi.org/10.1017/S0272263199003034

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty. C. (1995). Does textual enhancement prmote noticing? A think-aloud protocol analysis. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 183-216). University of Hawaii.

LaBrozzi, R. M. (2014). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research, 20*(1), 75-91. http://doi.org/ 10.1177/1362168814561903

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL student's reading comprehension and learning of passive voice. *Language*

Learning, 57(1), 87-118.
https://doi.org/10.1111/j.1467-9922.2007.00400.x

Leow, R. P. (1997). Attention, awareness and foreign language behavior. *Language Learning, 47*(3), 467-505.

Leow, R. P. (2001). Attention, awareness and foreign language behavior. *Language Learning, 51*(1), 13-155.
https://doi.org/10.1111/j.1467-1770.2001.tb00016.x

Leow, R. P., Egi, T., Nuevo, A. M., & Tsai, Y. C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13*(2), 93-108.

Loewen, S., & Inceoglu, S. (2016). The effectiveness of visual input enhancement on the noticing and L2 development of the Spanish past tense. *Studies in Second Language Learning and Teaching, 6*(1), 89-110.
http://doi.org/10.14746/ssllt.2016.6.1.5

Naseri, E., & Khodabandeh, F. (2019). Comparing the impact of audio-visual input enhancement on collocation learning in traditional and mobile learning contexts. *Applied Research on English Language , 8*(3), 383-422.
http://doi.org/10.22108/are.2019.115716.1434

Nassaji, H., & Fotos, S. (2014). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context.* Routledge.

Overstreet, M. H. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics, 2*, 229-258.

Philp, J. (2003). Constraints on noticing the gap. *Studies in Second Language Acquisition, 25*(1), 99-126.
https://doi.org/10.1017/S0272263103000044

Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10*(4), 277-303.
https://doi.org/10.1016/S1060-3743(01)00046-7

Robinson, P. (1995). Attention, memory, and the noticing hypothesis. *Language Learning, 45*(2), 83-331.
https://doi.org/10.1111/j.1467-1770.1995.tb00441.x

Rosa, E., & O'Neill, M. (1999). Explicitness, intake and the issue of awareness. *Studies in Second Language Acquisition, 21*(4), 511-556.
https://doi.org/10.1017/S0272263199004015

Russell, V. (2014). A closer look at the output hypothesis: The effect of pushed output on noticing and inductive learning of the Spanish future tense. *Foreign Language Annals, 47*(1), 25-47.
http://doi.org/10.1111/flan.12077

Sarkhosh, M., Taghipour, B., & Sarkhosh, H. (2013). Diffedrential effect of different textual enhancement formats on intake. *Procedia-Social and Behavioral Sciences, 70*(2), 544-559.
http://doi.org/10.1016/j.sbspro.2013.01.092

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.
https://doi.org/10.1093/applin/11.2.129

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press.

Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System, 37*(1), 124-135.
https://doi.org/10.1016/j.system.2008.06.005

Simard, D., & Wong, W. (2001). Alertness, orientation and detection. *Studies in Second Language Acquisition, 23*(1), 103-124.
https://doi.org/10.1017/S0272263101001048

Song, M. (2007). Getting learners' attention: Typographical input enhancement, output, and

their combined effects. *English Teaching, 62*(2), 193-215. https://doi.org/10.15858/engtea.62.2.200706.193

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Madden (Ed.), *Input in second language acquisition* (pp. 235-253). Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In B. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honor of H. Widdowson* (pp. 125-144). Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teachning and learning* (pp. 471-484). Erlbaum.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371-391. https://doi.org/10.1093/applin/16.3.371

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183-204. https://doi.org/10.1017/S0272263100012870

Uggen, M. S. (2012). Reinvestigating the noticing function of output. *Language Learning, 62*(2), 506-540. https://doi.org/10.1111/j.1467-9922.2012.00693.x

VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, commentary* (pp. 5-32). Erlbaum.

White, J. (1998). Getting the learner's attention. In C. J. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition classroom* (pp. 85-113). Cambridge University Press.

Winke, P. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data . *Studies in Second Language Acquisition, 35*(2), 323-352. https://doi.org/10.1017/S0272263112000903

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Application of Computational Linguistics to Predicting Language Proficiency Level of Persian Learners' Textbooks

Research Article
pp. 29-52

Masood Ghayoomi[1]

Received: 2020/08/15 Accepted: 2021/02/09

Abstract
One subfield of assessment of language proficiency is predicting language proficiency level. This research aims at proposing a computational linguistic model to predict language proficiency level and to explore the general properties of the levels. To this end, a corpus is developed from Persian learners' textbooks and statistical and linguistic features are extracted from this text corpus to train three classifiers as learners. The performance of the models vary based on the learning algorithm and the feature set(s) used for training the models. For evaluating the models, four standard metrics, namely accuracy, precision, recall, and F-measure were used. Based on the results, the model created by the Random Forest classifier performed the best when statistical features extracted from raw text is used. The Support Vector Machine classifier performed the best by using linguistic features extracted from the automatically annotated corpus. The results determine that enriching the model and providing various kinds of information do not guarantee that a classifier (learner) performs the best. To discover the latent teaching methodology of the textbooks, the general performance of the classifiers with respect to the language level and the linguistic knowledge used for creating the model are studied. Based on the obtained results, the amount of extracted features plays an important role in training a classifier. Furthermore, the average best performance of the classifiers is extending the linguistic knowledge from syntactic patterns at proficiency level A (beginner) to all linguistic information at levels B (intermediate) and C (advanced).

keywords: machine learning, classification, feature, computational cognitive model, Persian learner

[1] Assistant Professor, Faculty of Linguistics, Institute for the Humanities and Cultural Studies, Tehran, Iran; M.Ghayoomi@ihcs.ac.ir

DOI: 10.22051/LGHOR.2021.32656.1354

Introduction

Language proficiency assessment is a fundamental step within the language learning process. This task should be done precisely to be used in fundamental decisions such as demanding for work or obtaining the study permission. To reach the goal, various exams are compiled to precisely evaluate the language skills, including listening, speaking, reading, and writing. To this end, Test Of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS) as the two widespread, major, distinguished tests are designed to evaluate the proficiency level of English. The Common European Framework of Reference for Languages (CEFR; Council of Europe 2001) is a framework of reference to provide a transparent, coherent, and comprehensive basis for the language teaching syllabuses and guidelines. This framework is used in Europe and also in other continents. The application of using this framework caused to know 3 groups of language proficiency levels, namely beginner (A), intermediate (B), and advanced (C). The levels are extended to six-point scale known as the basic user (A1 and A2), the dependent user (B1, B3), and the proficient user (C1, and C2). CEFR guideline has provided a complete description for each language level and defined the properties of each level.

Language learning procedure has also been studied from cognitive linguistics perspective (Belkhir, 2020). Matlin (2005, p. 2) defined "cognition" as a mental activity with various cognitive processes. This statement determines how wide the cognition is. One property of this field is to translate the CEFR language levels' properties into linguistic features. These features can be transformed in such a way to be used by a computer to simulate the language learning process and create a model. It should be noted that, as Macwhinney (2010) stated, in computational modeling, the whole process of language learning cannot be simulated but partly.

In this paper, we aim to determine the language proficiency level of a given written text automatically by using computational linguistics methods. This approach explores the general content properties of Persian learners' textbooks. The outcome of this research can be used to discover the latent teaching methodology in the Persian learners' textbooks and to check their major and minor focuses on a linguistic knowledge. This knowledge can be utilized to increase the content quality of the textbooks and to have a uniformed distribution of the linguistic knowledge for all levels. One additional application of creating such a model is predicting the language proficiency level of essays in the second language and discovering the linguistic knowledge that the learner is master at it.

The structure of the paper is as follows: in Section 2, we briefly overview the previous studies on developing methods to simulate the language learning process and to measure the language proficiency levels. The theoretical framework of second or foreign language learning is discussed in Section 3. Section 4 describes our proposed algorithm to determine the language proficiency level of a written text. In Section 5, the obtained results are reported and discussed. And finally, in Section 6, we conclude the paper.

Previous Studies

Two concerns are found in the literature, when one studies readability assessment. One is the application of studying the readability of a text for different topics, and the other one is the methodology how to assess the readability of a text.

There are a large number of researches that have focused on readability of a

Persian text from different perspective, for instance studying readability and linguistic properties of fictions (Khademizadeh & Vaezi, 2020; Vaezi et al., 2016), readability of school textbooks (Ghaderi Moghaddam & Sobhaninejad, 2016; Khodadady & Mehrazmay, 2017; Nazari et al., 2016; Shekari & Najareyan, 2012), readability of pharmaceutical brochures (Zarea Gavgani et al, 2018), readability of patients (Ahmadzadeh et al., 2014), readability of translation studies (Maftoon & Daghigh, 2001), readability of accounting standards (Sarvi et al., 2019), and readability of health information (Zeinali et al., 2019). In these studies, the readability property of various domains is studied. However, the research by Mohammadi and Khaste (2020) on Persian is the only research that used machinery methods to assess the readability of an open domain text collected through crowd-sourcing. In the data collection process, 400 people participated to collect 12780 texts. The data is labeled as easy, medium, and hard. In their proposed model, 5 classifiers were used to predict the difficulty level of the texts. To train the classifiers, vectors which contained statistical features, such as average sentence length, average word length, and linguistic patterns based on *n*-grams ($1 \leq n \leq 5$) extracted from word forms along with their part-of-speech tags, were used

Although there are studies such as Doró (2011) who used statistical analyses to predict educational success through language proficiency, ten Bosch et al. (2009) who proposed a computational model for human cognitive process to acquire a language, or Matusevych et al. (2013) who proposed a model to study the impact of construction priming and statistical distribution on learning a second language, the current research aims at predicting language proficiency level of a written Persian text automatically.

Luo et al. (2008) studied oral language proficiency through signal processing. The shadowing method was used such that the learners ought to repeat the utterance of the instructor immediately. Pronunciation of learners belonging to lower proficiency level had delays and errors and it could not match with the instructor's production.

de Wet et al. (2009) proposed a model to evaluate oral language proficiency and listening comprehension. To this end, a spoken dialogue system was developed. To evaluate the language proficiency, speech rate, pronunciation goodness as well as repeat accuracy were calculated. According to the results, speech rate provided a fair indication of oral proficiency.

Crossley et al. (2011) studied the application of lexical indices to predict the proficiency level of the language learners. In their study, a set of 1000 writing samples in second language was collected from 100 learners. The texts were categorized into three groups, namely beginner, intermediate, and advance. Wide range of lexical knowledge was extracted from the data. The extracted features were used to train a classification model to predict the language proficiency level. According to the results, 70% of the texts were correctly classified.

Pilán, Alfter et al. (2016) proposed a classification model and used a machine learning method to determine the language proficiency level of an essay written by a human learner and to enhance the model by utilizing reading passages in the language learning textbooks. Experimental results of the proposed model showed that incorporating features from the latter dataset boost the performance of the model significantly.

Pilán, Volodina et al. (2016) aimed at proposing a model to predict the language proficiency of Swedish. To this end, they used a classification model. One

main issue they studied was the data sparsity problem and using the data in second language that belonged to another domain to overcome this problem. In the model, they used statistical as well as linguistic features, including morphological, syntactical, and semantic features. These features were used to train the Support Vector Machine (SVM) classifier (Boser et al., 1992).

Yang et al. (2016) proposed a classification model to predict language level proficiency. They used linguistic cognitive properties to the model to improve the performance of their model.

Monaghan et al. (2017) explored the relations between word frequency, language exposure, and bilingualism in a computational model of reading.

Jung et al. (2019) used computational tools to predict the second language writing proficiency based on the learners' texts. In this study, it was found that linguistic features associated with text length and lexical complexity were the most important predictive elements of the writing quality.

McLean et al. (2020) used statistical methods to predict reading proficiency of the second language. In their study, the correlation between vocabulary knowledge and reading proficiency were calculated.

Theoretical Frameworks

There are three general language acquisition theories. One is the *nativist* approach that is rooted at the Chomskyan attitude towards language learning (Chomsky, 1965; 1968; 1980). Rationalists who follow this approach provide a formal representation of the language and they believe that language is instinct (Pinker, 1994) and there is an innate language learning capability in the brain known as universal grammar that is given to a language learner for free and tuning is required to fully acquire the language.

The other one is the *emergentist* explanation approach that is more functional and usage-based (MacWhinney, 1999). This approach rooted at this idea that the language structure emerges from the language use and it is not instinct (Evans, 2014). This attitude is nothing but patterns as a sequence of meaningful linguistic symbols that are induced by a child. Therefore, all complex cognitive activities as an abstract knowledge are acquired to construct one's language.

The third one is the *cognitive* approach within the constructivism theory that was proposed by two theorists Jean Peaget and Lev Vygotsky in 1930 and 1934, respectively (Sulistyowati, 2019). Constructivism is "an approach to learning that holds that people actively construct or make their own knowledge and that reality is determined by the experiences of the learner" (Elliott et al., 2000, p. 256). Constructivist item-based learning is proven to be children's early linguistic competence (MacWhinney, 2005; Tomasello, 2000); therefore, the language development process is from specific to general and it is organized around a concrete linguistic schema. For instance, to learn words, children experience various and rich contexts to construct the knowledge received from different contexts.

Undoubtedly each child acquires a language as a mother tongue language. But due to the living situation, (s)he might acquire additional language after the native language, called the second language. The second language can have a function in the place (s)he lives, but it should bear in mind that life situation may vary and there is a possibility that the person might learn a language for a very limited use in a classroom. This language is known as the foreign language. Theoretically, there is no difference between the second and foreign language

methods as Ellis (1997, p. 3) stated that the "second [language] is not intended to contrast with 'foreign' [language]. Whether you are learning a language naturally as a result of living in a country where it is spoken, or learning it in a classroom through instruction, it is customary to speak generically of 'second' language acquisition."

There are methods and approaches to teach a second or foreign language. Methods refer to prescriptions for a teacher and a learner on how to act during the teaching or learning process. Djigunović and Krajnović (2005) have briefly described the methods used for teaching or learning a second or foreign language. These methods that have been used for a long period of time include a) grammar translation, b) direct method, c) audio-lingual method, and d) the cognitive code learning.

Approaches are the theories on the nature of language and language learning. In addition to the language acquisition theories that were briefly described above, there are other approaches for teaching a second or foreign language, including Communicative Language Teaching (CLT), Task-based Language Learning (TLL), and Computer Assisted Language Learning (CALL).

In 1970s, CLT influenced language teaching (Brumfit & Johnson, 1979) such that language could be learnt through interaction. In this teaching approach, the mere attention to language structure was replaced by focus on the meaning and conveying information to one another (Widdowson, 1978).

In 1980s, CLT became mature and gradually the 'communicative' activity was replaced by a 'task'. Prabhu (1987) proposed a teaching approach called TLL. In TLL, learners are expected to do meaningful tasks using the target language, and much attention has been paid to meaning. Skehan (1998) and Robinson (2001) believed that, in TLL, learners make generalizations based on the forms in an attention-driven perspective. This approach fits with the constructivist item-based learning where the language develops from specific to general.

The first attempts to use CALL back to 1960s in the PLATO project. The development of this learning approach continued through progress in technology and the advent of microcomputers in the late 1970s (Marty, 1981). Levy and Stockwell (2006, pp. 249, 253) explored that "there is no one single TLL methodology" and CALL is both "interdisciplinary and multidisciplinary" in orientation. This idea persuaded Thomas and Reinders (2010, pp. 4-5) to find synergies and the common ground between TLL and CALL approaches. They introduced micro and macro processes in language learning. The micro-level strategy involves "practical teaching, task framework, design and evaluation decisions based on particular learning contexts" and the macro-level strategy involves "analysis and integrated task, syllabus and curriculum design". TLL contains both micro and macro processes, while CALL contains only the micro process. This attitude towards language learning brought technology into account and the term Technology-Enhanced Language Learning (TELL) was used in 1990s instead of CALL (Bush & Terry, 1997). Internet and Web-based applications introduced another term called E-learning, which is defined by UNESCO as a type of learning by means of Internet or multimedia. One advantage of E-learning is that the learner is self-dependent which activates the cognitive motivation.

One of the tasks for second or foreign language learning is learning the required vocabulary to achieve fluency in a language. There are two learning strategies, namely incidental (implicit) learning, and intentional (explicit) learning

(Postman & Keppel, 1969). In incidental learning, the vocabularies are learnt without placing the focus on a specific word to be learned (Paribakht & Wesche, 1999), while in intentional learning, the learners are aware in advance what is going to be learned. Uchihara et al. (2019) found that exposure to the target vocabulary in a second language, which was originally studied by Ebbinghaus (1964), for a certain number of times affects the likelihood of the vocabulary to be learned. This achievement means that repetition has an impact on learning. In repeated exposure method, words are learned in the diversity of contexts; as a result, the learners do not learn the word forms such as the verbs at once but generalized knowledge is obtained at a later learning stage (Tomasello, 1992). The repeated exposure method causes learners to find linguistic patterns through statistical learning and analogy to create a more abstract knowledge about the language (Tomasello, 2006). This idea explores the importance of statistical information in the cognitive approach of language learning. The statistical learning attitude towards language learning is totally ignored by Chomskyan linguistics because "any account which assigns a fundamental role to segmentation, categorization, analogy, and generalization" is rejected as "mistaken in principle" (Chomsky, 1975). However, there are a number of researches on the cognitive approach of the language learning that determine how a learner is sensitive to the statistical structure of their linguistic input (Aslin et al., 1998; Gomez & Gerken, 1999; Newport & Aslin, 2000; Saffran et al., 1996). These findings have made progress to propose computational models to simulate the language learning process. The computational models of cognitive process deepen understanding of how induction methods are used to learn a language. Pinker (1996, p. 13) explored that a mathematical learning model contains 4 parts: 1) properties of the language within the scope of the learner's acquisition capability; 2) instances that the computational model uses to learn; 3) the learning algorithm; 4) the criteria to evaluate the model and to make it possible to conclude how well the proposed algorithm works.

Within the cognitive item-based repeated exposure framework and the interdisciplinary property of CALL, in this paper, we propose a computational model to predict language proficiency level of Persian learners' textbooks by utilizing the linguistic and statistical properties of the language to build a statistical model and study the underlying methodology used for compiling the textbooks.

Proposed Algorithm

Our proposed algorithm to determine the language proficiency level of a text has 4 components, according to Pinker (1996): a) the data to build the model; b) the feature selection component to extract the required knowledge from the data; c) statistical classifiers as learners to use the features and to predict the language proficiency level of a text; and d) evaluation. The components are described in the following.

Data

The data used in our model is a collection of texts in textbooks for teaching Persian to non-Persian speakers. Various Persian learning textbooks at different levels, including the Basic (A), Intermediate (B), and Advanced (C) levels, are available that are listed below:

- **Source 1**: "Teaching the Persian Language" in basic, intermediate, and advanced levels by Samareh (1989; 2005a; 2005b; 2005c);

- **Source 2**: "The Persian Lesson for Foreign Persian Learners" in basic level by Poornamdariyan (1994);
- **Source 3**: "Series of Teaching the Persian Language" in basic, intermediate, and advanced levels by Zarghamiyan (1998; 2001a; 2001b);
- **Source 4**: "General Persian" in basic level by SaffarMoghaddam (2003);
- **Source 5**: "The Persian Language" in basic, intermediate, and advanced levels by Saffar Moghaddam (2008a; 2008b; 2008c; 2008d);
- **Source 6**: "Let's Learn Persian" in basic, intermediate, and advanced levels by Ghaffari et al (2004).

To collect this data as the corpus to be used in our model to extract the required statistical information, two people typed the print version of the selected passages in the books and provided the electronic format of the texts as a corpus. The selected passages were complete and coherent texts, mostly from the reading comprehension sections. The texts did not have a dialogue format, and the texts in the exercises were not used. Table 1 summarizes the general statistical information extracted from the target sources. It has to be emphasized that we assumed that the defined levels of the textbooks are correct and we utilized the levels as indicated by the labels to represent the proficiency levels.

Table 1
Detailed Statistical Information of Target Sources

Source	Level	Sentences	Tokens	Types	Lemmas
Source1	A	149	1657	656	591
	B	359	5310	2414	2106
	C	552	11190	4730	4260
Source2	A	272	2957	1585	1453
Source3	A	130	967	411	393
	B	373	3579	1714	1604
	C	776	10306	5000	4543
Source4	A	221	2828	1649	1580
Source5	A	88	2926	926	551
	B	135	2849	1653	1585
	C	209	4662	2431	2317
Source6	A	143	1269	699	672
	B	379	5470	3217	3049
	C	1633	28857	11698	10445

Comparing the extracted statistics of texts belonging to different levels in Table 2 reveals that, as the level goes higher, the number of sentences, word forms, and lemmas to compile the texts increase. Although the number of texts for different levels is almost balanced, Level C contains a large number of sentences, word forms, and lemmas.

Table 2
Summary of Statistical Information of Target Sources

Level	Texts	Sentence	Tokens	Types	Lemmas
A	88	1004	10604	1968	1645
B	84	1247	17206	3499	2863
C	87	3171	55004	8767	7112

Feature Selection

Basic Information. The collected raw corpus has to be normalized and tokenized to acquire reliable results. Then, the data has to be analyzed linguistically. The linguistic information that is added to the data varies from phonological, to morphological, syntactical, and semantic information.

In the normalization process, character codes for the Arabic letters Yeh "ي" and Kāf "ك" were replaced with the equivalent Persian letters as "ی" and "ک". In the tokenization process, space was added between the letters, punctuation marks, and numbers as a word boundary to recognize each token (word). These two tasks were done automatically. Moreover, the extra white space was replaced by a pseudo-space to resolve the multi-unit token problem (Ghayoomi, 2018). This task was done semi-automatically.

Five linguistic analyses were run automatically on the data. The first one was assigning phonological patterns to each word form. To this end, we used the productive lexicon developed by Eslami et al. (2004). This word list contains lemmas and their phonological patterns. In our research, we drove the phonological pattern for each word form from the lemmas, and then assigned the pattern to the target words.

The two next analyses were lemmatization and Part-Of-Speech (POS) tagging. To these ends, we used the Marmot tool (Müller et al., 2013) for POS tagging and the Lemming tool (Müller et al., 2015) for lemmatization. These two toolkits were adapted for Persian by Ghayoomi (2019a) who compared the performance of different tools to lemmatize and POS tag Persian words. Based on the reported experimental results, the toolkits Marmot and Lemming outperformed the other toolkits. These results persuaded us to use these tools to annotate our corpus. To train the tools with the Persian data, we used the Bijankhan Corpus (Bijankhan, 2004) that has already been POS tagged semi-automatically and Ghayoomi (2019b) lemmatized the corpus semi-automatically. The tag set that was used in the Bijankhan Corpus is fine-grained and it contains 586 POS tags. Ghayoomi (2012) standardized the POS tags according to the Multi-Text East standard. In this standard, the length of each tag became fixed with respect to the main category of the tag; furthermore, specific information was defined in certain positions. The standardized data was used for training the tools.

The annotated data after lemmatization and POS tagging was the input to two other tools to do the syntactic analyses and to provide the parse tree of sentences. To this end, constituent parsing and dependency parsing of the data were performed. The toolkit used for constituency parsing was the statistical Stanford Parser (Klein & Manning, 2003) adapted to Persian by Ghayoomi (2013). To train this toolkit, the Persian constituency treebank developed by Ghayoomi (2012) was used. This treebank was developed within the linguistic framework of Head-driven Phrase Structure Grammar (Pollard & Sag, 1994). One property of this treebank is

that four types of head- dependent relations are defined in the tree analysis of each sentence and the trees are decorated with this additional information. The treebank contains 1024 sentences from Bijankhan Corpus. The dependency parser used in our research was the Mate Parser (Bohnet, 2009). The dependency parser was trained with the dependency treebank developed through conversion from the constituency treebank by Ghayoomi and Kuhn (2014). The dependency treebank contained 49 unique dependency relations. In the next step, the data standardization was required. When all linguistic analyses were done, all information was collected and standardized according to the CoNLL data format (Ghayoomi, 2020).

The CoNLL data format was proposed by the Conference on Natural Language Learning (CoNLL) in 2006. In this data format, each word appears in one row and the related information is defined in columns separated by a tab delimiter. The sentence border is determined by an empty line. Table 3 shows an example how the data from various sources is structured in 9 columns according to the CoNLL format.

Table 3

Sample of the Organized Data based on the CoNLL Format

Sent. ID	Word ID	Word Form	Phonological Pattern	Lemma	POS tag	Dep. Relation	Dep. Type	Constituency Tree
1	1	هادی	CVCV	هادی	Nasp---	5	NSUBJ	(ROOT (S (VPS (Nasp--- *)
1	2	کارمند	CVCCVCCV	کارمند	Ncsp--z	1	NN	(VPC (NPC (Ncsp----z- *)
1	3	اداره	CVCVCVCV	اداره	Ncspk-z	5	COPCOMP	(NPC (Ncspk---z- *)
1	4	پست	CVCC	پست	Ncsp---	3	NN	(Ncsp------ *)))
1	5	است	CVCC	بودن	Vpykshs----	0	ROOT	(Vpyk-shs------ *))
1	6	.	.	.	Oe	5	PUNC	(Oe *)))

In this table, the following information is available:

- the sentence ID in a text;
- the word ID in a sentence;
- the word form;
- the phonological pattern of the word form;
- the lemma of the word form;
- the POS tag of the word form in the sentence;
- the ID of the head word on which the dependency tree analysis depends;
- the type of the dependency relation between the head and the dependent words;
- the constituency tree analysis in the CoNLL 2011 format (Ghayoomi, 2020) in such a way that the nodes in the tree are drawn till the target word form is met.

According to the available information described above, the basic statistical information was extracted from the annotated data related to the words and the whole text. This information for each word contains:

- the number of characters in each word;
- the number of syllables in each word;
- the number of the CVCC phonological pattern in each word;
- the number of the CVC phonological pattern in each word;

- the number of the CV phonological pattern in each word;
- the number of constituency nodes in the tree analysis related to each word;
- the number of named entities.

Training Features. Based on the basic statistical information mentioned in the previous section and the annotated data, 37 features were extracted from the data and represented as a 37- dimension vector to train the supervised machine learning models. We ought to bear in mind that the labels assigned to the vectors were the original textbooks' proficiency levels.

The features were categorized into 5 groups. The statistical information in the vectors was extracted from both the raw corpus and the annotated corpus that contained phonological, morphological, syntactical, and semantic labels. In the vector representation, all features were normalized according to the word, sentence, or the text. Therefore, the relative frequency rather than the absolute frequency was used to construct the vector in order to reduce the negative impact of the text length.

Before describing the features, it is necessary to compare the data we utilized in our research with the data that Mohammadi and Khasteh (2020) used. In our study, we used various linguistic information in addition to statistical information extracted from the corpus. This property made the features much richer than the features used by Mohammadi and Khasteh (2020) who used merely statistical information of word forms and POS tags. Moreover, we used the content of textbooks that are used for teaching Persian to non-native speakers. This property of the corpus controlled the data accurately in terms of lexicon, syntactic and semantic complexities in the texts; while the dataset developed by Mohammadi and Khasteh (2020) through crowd-sourcing was labeled by using people's intuition without making any scientific justification to assign a difficulty label to a text. In addition, if their labeled data is accepted, no balance was found in their developed data for different levels (54% as an easy text, 32% as a medium text, and 14% as a hard text). The weak point of their data is the imbalanced data problem, because dominance on a specific level misleads classifiers. However, we used a sort of balanced data in terms of the number of documents to train the classifiers as reported in Table 2 (34% for the level A, 32% for the level B, and 34% for the level C).

Statistical Information from Raw Corpus. The statistical feature set extracted from the raw corpus contained the following information:

(a) the average word length;
(b) the average sentence length;
(c) lexical diversity which is the ratio of word types over word tokens;
(d) the ratio of word types from the first 100 word tokens of a text over word tokens in a text;
(e) the ratio of words (unigrams) with frequency 1 over word tokens in a text;
(f) the ratio of words with frequency 1 over word tokens with frequency above 1;
(g) the ratio of word bigrams (sequence of two words) with frequency 1 over the total number of word bigrams in a text;
(h) the ratio of word bigrams with frequency 1 over word bigrams with frequency above 1;
(i) the ratio of word trigrams (sequence of three words) with frequency 1 over the total number of word trigrams in a text;
(j) the ratio of word trigrams with frequency 1 over word trigrams with

frequency above 1.

Statistical Information from Phonological Annotation. The extracted statistical feature set based on phonological annotation of the corpus contained the following information:

(a) the average syllable of the words in a text;
(b) the ratio of one-syllable words over word tokens;
(c) the ratio of two-syllable words over word tokens;
(d) the ratio of three-syllable words and above over word tokens;
(e) the ratio of the CVCC syllable pattern over the total number of syllables;
(f) the ratio of the CVC syllable pattern over the total number of syllables;
(g) the ratio of the CV syllable pattern over the total number of syllables;
(h) the ratio of one syllable words in the first 150 word tokens of a text over the total number of syllables;
(i) the ratio of three-syllable words and above in the first 100 word tokens of a text over the total number of syllables.

Statistical Information from Morphological Annotation. The extracted statistical feature set based on morphological annotation of the corpus contained the following information:

(a) the ratio of word lemmas over word tokens;
(b) the ratio of word lemmas over word types;
(c) the ratio of word lemmas with frequency 1 over word tokens;
(d) the ratio of word lemmas with frequency 1 over word tokens with frequency above 1.

Statistical Information from Syntactic Annotation. The extracted statistical feature set based on syntactic annotation of the corpus contained the following information:

(a) the ratio of functional words over word tokens;
(b) the ratio of functional words with frequency 1 over word tokens;
(c) the ratio of functional words over content words;
(d) the ratio of content words over word tokens;
(e) the ratio of content words with frequency 1 over word tokens;
(f) the ratio of content words over functional words;
(g) the ratio of POS bigrams with frequency 1 over the total number of POS bigrams in a text;
(h) the ratio of POS bigrams with frequency above 1 over the total number of POS bigrams in a text;
(i) the ratio of POS trigrams with frequency 1 over the total number of POS trigrams in a text;
(j) the ratio of POS trigrams with frequency above 1 over the total number of POS trigrams in a text;
(k) the ratio of dependency relation types over the total number of dependency relations in a text;
(l) the ratio of clause dependency relations over the total number of dependency relations in a text;
(m) the ratio of the number of nodes in constituency constructions over word tokens in a text;

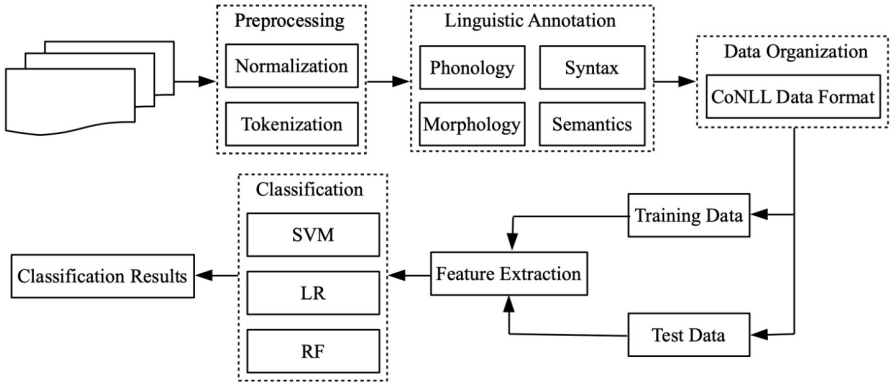
Statistical Information from Semantic Annotation. The extracted statistical feature set based on semantic annotation of the corpus contained the ratio of the named entities to word tokens in a text.

Our Proposed Algorithm

Figure (1) shows the architecture of our model. As can be seen in the figure, after data collection from Persian learners' textbooks, the corpus is pre-processed such that the texts are normalized to uniform the codes, and then the data is tokenized to identify each word. The cleaned data is linguistically annotated at the phonological, morphological, syntactical, and semantic levels. The annotated data from various levels should have a flat representation. To this end, we used the CoNLL data. A sample of this data structure is already shown in Table 3. We used Python programming language to develop our model.

Figure 1

Architecture of Our Proposed Model



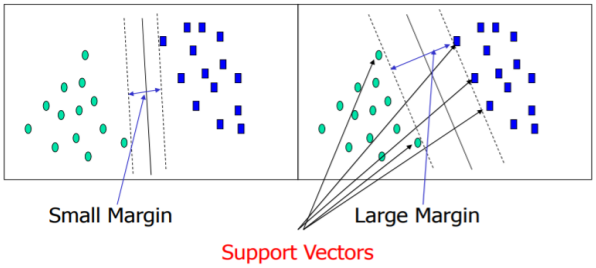
The prepared data that contains the labels of the proficiency levels is divided into two sets. The first division contains 90% of the data and it is used as training data to create the statistical model by a classifier. To this end, features are extracted from this data and they are represented as vectors to be used by the classifier. The second division of the data that contains 10% of the total data is used as test data to evaluate the performance of the classifier. For evaluation, we assume that the test data has no label and this data is given to the classifier to label. The output of the model on the test data and the original test data are compared to calculate the performance of the classifier.

Considering the nature of supervised learning techniques, the main two steps that have to be taken into the consideration are the feature engineering and the learning method. For the learning method, we utilize three algorithms that benefit from discriminative models, namely SVM (Boser et al., 1992), Logistic Regression (LR; Cramer, 2002), and Random Forest (RF; Breiman, 2001). One main property of discriminative models is that they use the inferred knowledge from a set of observed data. Although it seems that the three algorithms belong to a family, SVM works based on one single best margin with minimum risk of error, LR uses different weighted boundaries to make a near optimum decision, and RF uses stochastic discrimination over decision trees. Although deep neural networks have achieved state-of-the-art results in supervised learning, we cannot benefit from the deep neural approaches due to the small amount of the available data in our task. Therefore, three classifiers, namely SVM, LR, and RF, are used in our model to compare different sets of features and to study their impact on predicting language proficiency levels. Due to the shortage of amount data, no tuning for parameter

optimization is done and the default parameters in the classifiers are used. We use the scikit-learn[1] library in Python to call the classifiers into our code. The classifiers are briefly described in the followings.

SVM is a supervised machine learning method that uses the training data to build a model to assign a new instance to one or the other categories. To build the model, the data is represented in a vector space model. To make the distinctions, a hyperplane with maximum margin is used to create two subspaces. Figure (2) represents how the model decides between different possible hyperplanes based on their margin. The best hyperplane is the one with a large margin (James et al., 2013, pp. 337-342).

Figure 2
Support Vector Machine



LR, as another supervised machine leaning method utilized for classification, uses conditional probability assumptions that rely on the underlying data distribution represented as vectors. The probability score is a number between 0 to 1 (Indurkhya & Damerau, 2010, pp. 194-196).

RF (Breiman, 2001) consists of multiple random decision trees. While making the decision at each node tree, features are randomly selected to generate the best data split. In this classification model, feature selection has the most important contribution on the class probability. This means informative features confuse the model and they should be removed before fitting a classifier. However, redundant features reduce the importance of the features.

Evaluating the Algorithm

To evaluate the performance of our proposed algorithm, we use two evaluation metrics. One is calculating the accuracy of the model in Equation (1) for the total performance of the algorithm:

(1)

$$Accuracy = \frac{number\ of\ correctly\ predicted\ instances}{total\ \ number\ of\ instances}$$

The other evaluation metric is calculating F-measure in Equation (2) that is a harmonic mean of precision and recall proposed by van Rijsbergen (1979):

(2)

$$F - measur = \frac{(1+\beta) \times P \times R}{P + R}$$

[1] https://scikit-learn.org/stable/

where P is precision, R is recall, and β is a weighting parameter. If $\beta > 1$, more weight is assigned to recall, and in case $\beta < 1$, more weight is assigned to precision. If $\beta = 1$, precision and recall are considered equally. Equations (3) and (4) compute precision and recall, respectively:

(3)

$$P = \frac{number\ of\ correctly\ predicted\ instances\ for\ level\ X}{number\ of\ predicted\ instances\ for\ level\ X}$$

(4)

$$r = \frac{number\ of\ correctly\ predicted\ instances\ for\ level\ X}{number\ of\ instances\ for\ level\ X\ of\ gold\ data}$$

where X is one of the proficiency levels A, B, or C. After calculating precision and recall for each language level, we calculate the average precision and recall for all of the levels.

Since the total amount of data that we use in our research is not much, we use 10-fold cross validation method to evaluate our model; that is, the total amount of data is divided into 10 folds and in each round of experiments, one fold (10% of the data) is considered as the test data, and the rest (90% of the data) as the training data. The average of the obtained results, for both accuracy and F-measure, can be considered as the performance of the model.

Results and Discussions

This research aims at proposing a computational linguistic model to predict language proficiency level and to explore the general properties of the levels. To reach the goal, we configured different feature sets defined in Section 4.2.2 to train the classifiers for different levels. To make the comparison possible and to represent the superiority of the rich features to train a model, we require a baseline. To this end, we use Term Frequency-Inverse Document Frequency (TF-IDF) which represents how relevant a word to a document is in a collection of documents (Salton et al, 1975).

We proposed learning scenarios with different feature sets. The first learning scenario contained all available features that are the mixture of both statistical and linguistic features. The second and third learning scenarios contained statistical information and the combination of all linguistic information respectively. The fourth to seventh learning scenarios contained individual linguistic information to build separate models, including phonetics, morphology, syntax and semantics. After training the classifiers, the test data was given to the classifiers.

The confusion matrix of the labeled data for different classifiers belonging to different levels is reported in Table 4. In this table, the predicted labels of each classifier trained with different feature sets are compared to the gold data of a specific language level. As an example, the first row of Table 4 should be read in such a way that the gold data has 88 texts which belong to the proficiency level A. If the TF-IDF feature set is used, the SMV classifier assigns the predicted label C to all data. If the Statistical and Linguistic Features are used, the SVM classifier predicts the level A for 20 texts, the level B for 42 texts, and the level C for 26 texts out of 88 texts. As a result, 20 texts are labeled correctly, and the rest of 68 texts are labeled incorrectly. Using different feature sets causes the classifiers to behave differently.

Table 4
Confusion Matrix for Labeling of Classifiers for All Levels

Classifier	Level	Gold	TF-IDF			Statistics, Linguistics			Statistics			Linguistics			Phonetics			Morphology			Syntax			Semantics		
			A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
SVM	A	88	0	0	88	20	42	26	28	31	29	31	48	9	16	65	7	24	31	33	40	32	16	36	22	30
	B	84	0	0	84	11	14	59	7	13	64	9	35	40	11	34	39	22	25	37	16	19	49	27	14	43
	C	87	0	0	87	14	5	68	12	10	65	1	6	80	3	2	82	9	10	68	2	2	83	16	9	62
LR	A	88	46	0	42	65	20	3	66	17	5	62	22	4	71	14	3	61	2	25	59	23	6	55	0	33
	B	84	43	0	41	21	44	19	22	43	19	21	50	13	35	20	29	60	1	23	23	47	14	33	1	50
	C	87	67	0	20	2	14	71	1	13	73	5	11	71	6	5	76	24	3	60	3	16	68	22	6	59
RF	A	88	37	0	51	66	19	3	66	17	5	69	16	3	68	17	3	37	27	24	67	16	5	53	9	26
	B	84	35	0	49	29	41	14	21	44	19	30	40	14	34	35	15	22	43	19	31	38	15	36	5	43
	C	87	58	0	29	1	9	77	1	11	75	3	9	75	3	8	76	9	17	61	0	18	69	21	12	54

Based on the number of labeled data in Table 4 for different classifiers and Equations (1) to (4), we evaluated the performance of the classifiers, using accuracy and F1 metrics. Table 5 reports the average performance of the models using 10-fold cross-validation method. In our experiments, we used TF-IDF as the basic feature set to compare the performance of the models that used different feature sets. As it is obvious from the results, the malperformance of the classifiers was obtained when trained with TF-IDF feature set. In general, the SVM classifier created the worst model, and the RF classifier outperformed the other models using features other than TF-IDF. This achievement determines that previous words (the history) play a very important role in the learning process. However, the SVM classifier in comparison to the LR and RF classifiers performed the best when using TF-IDF features set. The result shows that, for this classifier, history has a negative impact on predicting labels.

Comparing each classifier based on the feature sets that were used for training, the model created by SVM and LR classifiers performed the highest when a combination of linguistic information was used. Additionally, RF performed the best using statistical information. However, RF performed slightly worse than the model using either combination of linguistics information or combination of statistical and linguistic. This result determines that additional information could not improve the classifier's performance.

We further evaluated the performance of the classifiers using different features for each proficiency level. The results are reported in Table 6. While the SVM model performed the best when linguistic information was used in Table 5, the level C obtained the highest results among the three levels in Table 6. We further observed that using syntactic information for the level A in the SVM model outperformed using other feature sets at this level. The SVM model that used all linguistic information performed well for the level B; and using the phonetic information to build the SVM model caused to perform the best for the level C.

Comparing the performance of the LR models, the model utilized the linguistic information to build the model obtained the highest results in Table 5, and

the level C for this feature set obtained the highest result among the three levels in Table 6. We further observed that using statistical information for the level A in the LR model outperformed using other feature sets at this level. Additionally, the LR model that utilized all linguistic information performed the best for the levels B and C.

Comparing the performance of the RF models for different levels, the utilized statistical information in Table 5 that obtained the highest results performed the highest for the level C in Table 6. We further observed in Table 6 that using statistical information for the levels A and B in the RF model outperformed the models that utilized other features. Moreover, the RF model that used both statistical and linguistic information performed the best for the level C.

Table 5
Performance of the Classifiers Using Different Feature

Classifier	Evaluation Metric	TF-IDF	Statistics Linguistics	Statistics	Linguistics	Phonetic	Morphology	Syntax	Semantic
SVM	Accuracy	33.54	39.35	40.85	56.32	50.94	45.12	54.82	43.20
	F1	16.31	37.57	40.46	57.95	51.33	44.91	54.99	41.94
	Recall	33.33	39.60	41.24	56.69	51.32	46.00	55.37	42.16
	Precision	11.18	35.75	39.70	59.26	51.34	43.86	54.61	41.73
LR	Accuracy	25.46	69.51	70.28	70.62	64.48	47.09	67.17	44.42
	F1	13.41	69.56	70.45	70.80	64.31	40.01	68.06	37.45
	Recall	33.33	69.58	70.63	70.92	64.53	47.60	68.26	44.92
	Precision	8.49	69.55	70.27	70.68	64.09	34.50	67.88	32.12
RF	Accuracy	25.46	71.05	71.42	71.05	69.08	54.37	67.20	43.28
	F1	13.41	70.57	71.96	71.14	68.44	55.29	66.50	38.33
	Recall	33.33	71.06	72.23	71.62	68.95	55.91	66.91	42.48
	Precision	8.49	70.07	71.68	70.66	67.94	54.68	66.09	34.92

Table 6
F-measure of the Classifiers Using Different Features for each Individual Label

Classifier	TF-IDF			Statistics, Linguistics			Statistics			Linguistics			Phonetic			Morphology			Syntax			Semantic		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
SVM	0	0	50.29	30.08	19.31	56.67	41.48	18.84	53.06	48.06	40.46	74.07	27.12	36.76	76.28	33.57	33.33	60.44	54.79	27.74	70.64	43.11	21.71	55.86
LR	37.70	0	21.05	73.86	54.32	78.89	74.58	54.78	79.35	70.45	59.88	81.14	71.00	32.52	77.95	52.36	2.22	61.54	68.21	55.29	77.71	55.56	2.20	51.53
RF	33.94	0	26.85	71.74	53.59	85.08	75.00	56.41	80.65	72.63	53.69	83.80	70.47	48.61	83.98	47.44	50.29	63.87	72.04	48.72	78.41	53.54	9.09	51.43
Average	23.88	0	32.73	58.56	42.41	73.55	63.69	43.34	71.02	63.71	51.34	79.67	56.2	39.3	79.4	44.46	28.61	61.95	65.01	43.92	75.59	50.74	11	52.94

Additionally, we investigated which feature set has impact on the performance of a classifier with respect to the language proficiency level. In this analysis, we calculated the average performance of the classifiers to focus on the latent teaching methodology used for compiling the textbooks. The results are reported in the last row of Table 6. The average best performance of the learners for the level A is achieved by using syntactic information; i.e. the syntactic information is the main focuses of the textbooks at this level to make the learners aware of the basic constructions. The average best performance of the learners for the levels B and C is achieved by using linguistic information. This result determines that a great amount of attention is given to linguistic information at levels B and C. Among the linguistic information, syntactic and phonetic information play the most important role for levels B and C, respectively. One reason is extracting various features for these two linguistic components from the annotated data; i.e., the higher the number of features to be extracted, the better a classifier learns about the properties of the language.

Concluding Remarks

In this paper, we proposed a computational linguistic model to detect the language proficiency level of a given text and to label it automatically by using machine learning methods. To reach the goal, we defined sets of features including statistical and linguistics. The linguistic features contained phonological, morphological, syntactic, and semantic features. Additionally, we used TF-IDF feature set as the baseline to compare the performance of the models. Our proposed models that used statistical and/or linguistic features outperformed the baseline. The features were extracted from a corpus developed from 6 Persian learners' textbooks that belonged to the beginner, intermediate, and advanced levels. The collected data was divided into subsets to train and test the classifiers. Based on the results, the

model created by the RF classifier performed the best using statistical features. This determined that enriching the model and providing more information does not guarantee to achieve the best performance. But this was not a global finding because it totally depends on the learning algorithm of the classifier, because the linguistic information caused the SVM classifier to perform the best.

We studied the performance of the classifiers with respect to the language proficiency level and the linguistic knowledge used to create the model. One property of the texts at the level A was paying much attention to the syntactic constructions. The general property of the texts at the levels B and C was using all linguistic information to compile the textbooks.

The outcome of this research can be used to check major and minor focus of the Persian learners' textbooks on linguistic knowledge and to increase the quality of the textbooks by utilizing uniformed distribution of linguistic knowledge for all levels. A sample for minor focus of the textbooks is the low performance of the classifiers for the level B when the semantic feature is utilized and the classifiers of this level had a malperformance compared to the levels A and C.

Acknowledgement

This research is funded by Iran National Science Foundation (INSF) for the research proposal number 97000696.

References

Ahmadzadeh, K., Khosravi, A., Arastoopoor, S., & Tahmasebi, R. (2014). Assessing the readability of patient education materials about diabetes available in Shiraz Health Centers. *Iranian Journal of Medical Education, 14*(8), 661-667. http://ijme.mui.ac.ir/article-1-3157-en.pdf

Aslin, R., Saffran, J., & Newport, E. (1998). Computation of conditional probability statistics by 8-month old infants. *Psychological Science*, *9*, 321-324. https://doi.org/10.1111/1467-9280.00063

Belkhir, S. (2020). Cognition and language learning: An introduction. In S. Belkhir (Ed.), *Cognition and language learning* (pp. 1-12). Cambridge Scholars Publishing.

Bijankhan, M. (2004). The role of corpora in writing a grammar: Introducing a software. *Journal of Linguistics*, *19*(2), 48-67.

Bohnet, (2009). Efficient parsing of syntactic and semantic dependency structures. In *Proceedings of the 13th conference on computational natural language learning: Shared task* (pp. 67-72). Association for Computational Linguistics. https://www.aclweb.org/anthology/W09-1210.pdf

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the 5th annual workshop on computational learning theory* (pp. 144-152). https://doi.org/10.1145/130385.130401

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Brumfit, C., & Johnson, K. (1979). *The communicative approach to language teaching*, Oxford University Press.

Bush, M., & Terry, R. (1997). *Technology-enhanced language learning*, National Textbook Company.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT Press.

Chomsky, N. (1968). *Language and Mind*. Harcourt Brace Jovanovich.

Chomsky, N. (1975). *Reflections on language*. Pantheon Books.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*, 1-61. https://doi.org/10.1017/S0140525X00001515

Cramer, J. S. (2002). *The origins of logistic regression*. Technical Report (pp. 167-178). Tinbergen Institute. https://papers.tinbergen.nl/02119.pdf

Crossley, S. A., Salsbury, T., & McNamara, D. S. (2011). Predicting the proficiency level of language learners using lexical indices. *Language Testing*, *29*(2), 240-260. https://doi.org/10.1177/0265532211419331

de Wet, F., Van Der Walt, C., & Niesler, T. R. (2009). Automatic assessment of oral language proficiency and listening comprehension. *Speech Communication*, *52*, 864-874. https://doi.org/10.1016/j.specom.2009.03.002

Djigunović, J. M., & Krajnović, M. M. (2005). Language teaching methodology and second language acquisition. In V. Muhvic-Dimanovski & L. Socanac (Eds.), *Encyclopedia of life support systems*, (pp. 394-417). Eolss Publishers Co. Ltd.

Doró, K. (2011). English language proficiency and the prediction of academic success of first-year students of English. *UPRT 2010: Empirical studies in English applied linguistics* (pp. 173-186). Lingua Franca Csoport. http://publicatio.bibl.uszeged.hu/11049/1/Doro%202011%20Language%20proficiency%20and%20academic%20success.pdf

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. (H. A. Ruger & C. E. Bussenius, Trans.). Dover Publications. (Original work published 1885). https://doi.org/10.5214/ans.0972.7531.200408

Elliott, S. N., Kratochwill, T. R., Littlefield, C., J., & Travers, J. (2000). *Educational psychology: Effective teaching, effective learning (3rd Ed.)*. McGraw-Hill College.

Ellis, R. (1997). *Second language acquisition*. Oxford University Press.

Eslami, M., Mosavi Atashgah, M., Alizadeh Lamjiri, S., & Zandi, T. (2004). Persian productive lexicon. In *Proceedings of the 1st workshop on the Persian language and computer*, University of Tehran.

Evans, V. (2014). *The language myth: Why language is not an instinct*. Cambridge University Press.

Ghaderi Moghaddam, M. E., & Sobhaninejad, M. (2016). Validation methods to measure textbooks readability. *Research in Curriculum Planning*, *13*(21), 44-55.

Ghaffari, M., Mahmoodi Bakhtiyari, B., & Zolfaghari, H. (2004). *Let's learn Persian* (Volumes 1-3). Madreseh Publication.
https://jsr-e.khuisf.ac.ir/article_534415_65a3945c9994bc90c81c23ab0eacfaf7.pdf?lang=en

Ghayoomi, M. (2012). Bootstrapping the development of an HPSG-based treebank for Persian. *Linguistic Issues in Language Technology*, *7*(1).

Ghayoomi, M. (2013). Word clustering for Persian statistical parsing. In H. Isahara, & K. Kanzaki, (Eds.), *Advances in natural language processing*, (pp. 126-137). Springer.
https://doi.org/10.1007/978-3-642-33983-7_13

Ghayoomi, M. (2018). The problem of multi-words in syntactic processing of Persian. In *Proceedings of the fourth nation conference on computational linguistics* (pp. 11-40). Neviseh Parsi Publications.

Ghayoomi, M. (2019a). *Studying issues for automatic processing of the Persian language on lemmatization, part-of-speech tagging, and parsing: Developing a software using machine learning methods*. Technical Report. Tehran, Iran.

Ghayoomi, M. (2019b). Transition from rule-based to statistical lemmatization in Persian. In *Proceedings of the 5th national conference on computational linguistics* (pp. 57-86). Neveeseh Parsi Publications.

Ghayoomi, M. (2020). Structuring multilayer linguistic analyses in linguistic corpora. In F. Ghatreh & Sh. Modarres Khiabani, (Eds.), *Word by word of life: Festschrift for professor Vida Shaghaghi* (pp. 287-312). Neveeseh Parsi Publications.

Ghayoomi, M., & Kuhn, J. (2014). Converting an HPSG-based treebank into its parallel dependency-based Treebank. In *Proceedings of the 9th international conference on language resources and evaluation* (pp. 802-809). Reykjavik, Iceland.
http://www.lrec-conf.org/proceedings/lrec2014/pdf/441_Paper.pdf

Gomez, R., & Gerken, L. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109-135.
https://doi.org/10.1016/S0010-0277(99)00003-7

Indurkhya, N., & Damerau, F. J. (2010). *Handbook of natural language processing*. Chapman & Hall.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer.

Jung, Y. J., Crossley, S., & McNamara, D. (2019). Predicting second language writing proficiency in learner texts using computational tools. *Journal of Asia TEFL*, *16*(1), 37-52.
https://doi.org/10.18823/asiatefl.2019.16.1.3.37

Khademizadeh, S., & Vaezi, M. R. (2020). Evaluating readability of Persian fictions selected by flying Turtle the Iranian award. *Publishing Research Quarterly*, *36***,** 116-128.
https://doi.org/10.1007/s12109-019-09705-0

Khodadady, E., & Mehrazmay, R. (2017). Evaluating two high intermediate EFL and ESL textbooks: A comparative study based on readability indices. *Sociology International Journal*, *1*(3), 93-102.
https://doi.org/10.15406/SIJ.2017.01.00016

Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. In *Proceedings of the 41st meeting of the association for computational linguistics* (pp. 423-430).
https://doi.org/10.3115/1075096.1075150

Levy, M., & Stockwell, G. (2006). *CALL dimensions: Options and issues in computer*

assisted language learning. Lawrence Erlbaum Associates.

Luo, D., Minematsu, N., Yamauchi, Y., & Hirose, K. (2008). Automatic assessment of language proficiency through shadowing. In *Proceedings of 6th international symposium on Chinese spoken language processing* (pp. 41-44). https://doi.org/10.1109/CHINSL.2008.ECP.22

MacWhinney, B. (1999). *The emergence of language*. Lawrence Erlbaum Associates.

MacWhinney, B. (2005). Item-based constructions and the logical problem. In *Proceedings of the workshop on psychocomputational models of human language acquisition* (pp. 53-68). Ann Arbor, Michigan. https://doi.org/10.3115/1654524.1654531

MacWhinney, B. (2010). Computational models of child language learning: An introduction. *Journal of Child Language*, *37*(3), 477-485. https://doi.org/10.1017/S0305000910000139

Maftoon, P., & Daghigh, M. (2001). Metric of determining readability of translated texts from English into Persian. *Humanities Bulletin*, *29*, 61-80. https://www.sid.ir/fa/journal/ViewPaper.aspx?id=27487

Marty, F. (1981). Reflections on the use of computers in second language acquisition. *System, 9*(2), 85-98. https://eric.ed.gov/?id=ED218932

Matlin, M. W. (2005). *Cognition*. John Wiley and Sons.

Matusevych, Y., Alishahi, A., & Backus, A. (2013). Computational simulations of second language construction learning. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 47-56). Sofia, Bulgaria. Association for Computational Linguistics. https://www.aclweb.org/anthology/W13-2606.pdf

McLean, S., Stewart, J., & Batty, A. O. (2020). Predicting L2 reading proficiency with modalities of vocabulary knowledge: A bootstrapping approach. *Language Testing*, *37*(3), 389-411. https://doi.org/10.1177/0265532219898380

Mohammadi, H., & Khasteh, S. H. (2020). A machine learning approach to Persian text readability assessment using a crowd-sourced dataset. In *Proceedings of the 28th Iranian conference on electrical engineering*, University of Tabriz. https://doi.org/10.1109/ICEE50131.2020.9260933

Monaghan, P., Chang, Y. N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language*, *93*, 1-21. https://doi.org/10.1016/j.jml.2016.08.003

Müller, T., Cotterell, R., Fraser, A., & Schütze, H. (2015). Joint lemmatization and morphological tagging with lemming. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 2268-2274). Lisbon, Portugal. Association for Computational Linguistics. https://www.aclweb.org/anthology/D15-1272.pdf

Müller, T., Schmid, H., & Schütze, H. (2013). Efficient higher-order CRFs for morphological tagging. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 322-332). Seattle, Washington, USA. Association for Computational Linguistics. https://www.aclweb.org/anthology/D13-1032.pdf

Nazari, F., Farhadpour, M. R., & Soleymani, E. (2016). Measure the readability of the Persian text of the 'Lets know More' section of the Quran book for the grades two, three, and four of elementary school based on the Flash-Diani and Galing-Diani formulas. *Quarterly Journal of Knowledge and Information Management*, *3*(3), 85-92. http://lib.journals.pnu.ac.ir/article_4415_f2b05f84f03592edc72327a8a72ec55b.pdf?lang=en

Newport, E., & Aslin, R. (2000). Innately constrained learning: Blending old and new

approaches to language acquisition. In S. Howell, S. Fish, & T. Keith-Lucas, (Eds.), *Proceedings of the 24th annual Boston University conference on language development*, Somerville, MA. Cascadilla Press.

Paribakht, T., & Wesche, M. (1999). Reading and 'incidental' L2 vocabulary acquisition: An introspective study of lexical referencing. *Studies in Second Language Acquisition*, *21*(1), 195-224. https://doi.org/10.1017/S027226319900203X

Pilán, I., Alfter, D., & Volodina, E. (2016). Coursebook texts as a helping hand for classifying linguistic complexity in language learners' writings. In *Proceedings of the workshop on computational linguistics for linguistic complexity,* (pp. 120-126). Osaka, Japan. https://www.aclweb.org/anthology/W16-4114.pdf

Pilán, I., Volodina, E., & Zesch, T (2016). Predicting proficiency levels in learner writings by transferring a linguistic complexity model from expert-written coursebooks. In *Proceedings of the 26th international conference on computational linguistics: Technical papers* (pp. 2101-2111). Osaka, Japan. https://www.aclweb.org/anthology/C16-1198.pdf

Pinker, S. (1994). *The language instinct*. William Morrow and Company.

Pinker, S. (1996). *Language learnability and language development*. Harvard University Press.

Pollard, C. J., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. University of Chicago Press.

Poornamdariyan, T. (1994). *The Persian lesson for foreign Persian learners (For beginners)*. Institute for Humanities and Cultural Studies Publications.

Postman, L., & Keppel, G. (1969). *Verbal learning and memory*. Penguin Books.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.

Robinson, P. (2001). Task complexity, cognitive load, and syllabus design. In P. Robinson, (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge University Press.

Saffar Moghaddam, A. (2003). *General Persian: Basic constructions*. Council of Extending Persian Language and Linguistics at the Institute for Humanities and Cultural Studies.

Saffar Moghaddam, A. (2008a). *The Persian language* (Vol. 1). Council of Extending Persian Language and Linguistics at the Institute for Humanities and Cultural Studies.

Saffar Moghaddam, A. (2008b). *The Persian language* (Vol. 2). Council of Extending Persian Language and Linguistics at the Institute for Humanities and Cultural Studies.

Saffar Moghaddam, A. (2008c). *The Persian language*. (Vol. 3). Council of Extending Persian Language and Linguistics at the Institute for Humanities and Cultural Studies.

Saffar Moghaddam, A. (2008d). *The Persian language*. (Vol. 4). Council of Extending Persian Language and Linguistics at the Institute for Humanities and Cultural Studies.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928. https://doi.org/10.1126/science.274.5294.1926

Salton, G. M., Andrew W., & Chung-Shu Y. (1975). A vector space model for automatic indexing. *Communications of the ACM*, *18*(11), 613-620. https://doi.org/10.1145/361219.361220

Samareh, Y. (1989). *Teaching the Persian language* (Vol. 1). Alhoda International Publications.

Samareh, Y. (2005a). *Teaching the Persian language* (Vol. 2). Alhoda International Publications.

Samareh, Y. (2005b). *Teaching the Persian language*. (Vol. 3). Alhoda International Publications.

Samareh, Y. (2005c). *Teaching the Persian language*. (Vol. 4). Alhoda International Publications.

Sarvi, A., Talebnia, G., Pourzamani, Z., & Jahanshad, A. (2019). Assessment readability and understandability of accounting standards by accountants and auditors using Flesch and Cloze Indexes. *Applied Research in Financial Reporting***,** *7*(2), 241-274. http://www.arfr.ir/article_85308_8ee110e57414180e4fc5eec833f18000.pdf?lang=en

Shekari, A., & Najareyan, Z. (2012). A study of the readability of Hedyehaye Asemani textbooks for grades four and five based on Gunning scale. *Research in Curriculum Planning*, *9*(6), 71-79. http://jsr-e.khuisf.ac.ir/article_534233_1f574dc5383e52c94da235658f255a2a.pdf

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

Sulistyowati, T. (2019). Bottom-up and top-down listening progress within cognitive constructivist learning theory. *Prominent Journal of English Studies*, *2*(1), 92-100. https://doi.org/10.24176/pro.v2i1.2962

ten Bosch, L., Boves, L., Van Hamme, H., & Moore, R. K. (2009). A computational model of language acquisition: The emergence of words. *Fundamenta Informaticae, 90*(3), 229-249. https://doi.org/10.3233/FI-2009-0016

Thomas, M., & Reinders, H. (2010). Deconstructing tasks and technology. In M. Thomas & H. Reinders, (Eds.) *Task-based language learning and teaching with technology* (pp. 1-13). Continuum International Publishing Group.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge University Press.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Early language development*, *4*(4), 156-163. https://doi.org/10.1016/S1364-6613(00)01462-5

Tomasello, M. (2006). Acquiring linguistic constructions. In D. Kuhn & R. Siegler (Eds.) *Handbook of child psychology* (pp. 255-298). Wiley. https://doi.org/10.1002/9780470147658.chpsy0206

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language learning*, *69*(3), 559-599. https://doi.org/10.1111/lang.12343

Vaezi, M. R., Kokabi, M., & Ebrahimi, M. (2016). Investigation of the readability level of authored fiction books, selected by Children's Book Council of Iran. *Research on Information Science & Public Libraries*, *21*(4), 629-649. http://publij.ir/article-1-1085-fa.pdf

van Rijsbergen, C. J. (1979). *Information retrieval*, 2nd ed. Butterworth-Heinemann.

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford University Press.

Yang, Y., Yu, W., & Lim, H. (2016). Predicting second language proficiency level using linguistic cognitive task and machine learning techniques. *Wireless Pers Commun, 86*, 271-285. https://doi.org/10.1007/s11277-015-3062-2

Zarea Gavgani V., Mirzadeh-Qasabeh, S., Hanaee, J., & Hamishehkar, H. (2018). Calculating reading ease score of patient package inserts in Iran. *Drug Healthc Patient Safety*, *19*(10), 9-19. https://doi.org/10.2147/DHPS.S150428

Zarghamiyan, M. (1998). *Series of teaching the Persian language (From Beginner to Advanced)* (Vol. 1). Council of Extending Persian Language and Linguistics.

Zarghamiyan, M. (2001a). *Series of teaching the Persian language (From Beginner to Advanced)* (Vol. 2). Council of Extending Persian Language and Linguistics.

Zarghamiyan, M. (2001b). *Series of teaching the Persian language (From Beginner to Advanced)* (Vol. 3). Council of Extending Persian Language and Linguistics.

Zeinali, V., Haghparast, A., Damerchilou, M., & Vazifehshenas, N. (2019). Quality and

readability of online health information produced by the Ministry of Health and Medical Education of Iran. *Journal of Health Administration*, *21*(74), 65-74. http://jha.iums.ac.ir/article-1-2798-en.pdf

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Implementing Alternative Assessment to Foster Organizational and Pragmatic Competencies in Oral Tasks: An EFL Context Study

Research Article
pp. 53-71

Fatemeh Hemmati*[1]
Mahboobeh Mortazavi[2]
Rouhollah Sharifi[3]

Received: 2020/09/02 Accepted: 2021/02/14

Abstract
The present study set out to inspect how organizational and pragmatic knowledge can be functions of the assessment techniques practices implemented in the classroom. To this aim, the effects of teacher, peer, and self-assessment techniques that focused on providing feedback on pragmatic and organizational knowledge were compared. A total number of 98 female Iranian learners studying at pre-intermediate level in a language institute participated in the study. They were randomly assigned to one of the three groups of self-assessment (SA) (N = 32), peer-assessment (PA) (N = 33), and teacher-assessment (TA) (N = 33). Pragmatic competence was gauged using Discourse Role Play Talks (DRPTs) and organizational knowledge was assessed by a scale covering the grammatical and textual knowledge of the participants. The results of the analysis of one-way ANOVA test indicated that self-assessment followed by peer assessment had an advantage over the teacher assessment technique in promoting both pragmatic and organizational competence. In general, the obtained results yielded support for the employment of alternative methods of assessment as pedagogical tools to foster language competence. Implications for EFL pedagogy are discussed.

keywords: self-assessment, peer-assessment, teacher-assessment, pragmatic knowledge, organizational knowledge

*Corresponding Author
[1] Associate professor, Department of English Language Teaching, Payame Noor University (PNU), P.O.Box 19395-4697, Tehran, Iran, f.hemmati@jonoub.tpnu.ac.ir
[2] PhD, Department of English Language Teaching, Payame Noor University, (PNU), P.O.Box 19395-4697,Tehran, Iran, Mortazavi.0m@gmail.com
[3] Assistant Professor, Department of English Language Teaching, Payame Noor University (PNU), P.O.Box 19395-4697, Tehran, Iran, rsharifi@pnu.ac.ir

DOI: 10.22051/lghor.2021.32932.1361

Introduction

In the past decades, language researchers have come to acknowledge the role of competences other than linguistic competence in one's ability to become involved in communication and have even emphasized the explicit instruction of pragmatic competence (e.g., Ifantidou, 2013; Ishihara & Cohen, 2014; Sykes, 2013; Taguchi & Roever, 2017). Theoreticians and scholars (e.g., Bachman, 1990; Bachman & Palmer, 1996, 2010; Canale & Swain, 1980; Hymes, 1972) have asserted that the ability to use language comprises not only the ability to produce structurally accurate utterances and the possession of linguistic knowledge but also the awareness of the appropriateness of the language attuned to the demands of the specific situations, with the latter being called pragmatic competence. According to Bachman's (1990) model of communicative language ability, which was later on refined by Bachman and Palmer (1996, 2010), language competence comprises of two main competencies of organizational knowledge and pragmatic knowledge, the former being further classified into grammatical and textual knowledge and the latter embedding sociolinguistic and functional knowledge or illocutionary knowledge referring to the intention behind the sentences. In other words, pragmatic competence is the ability of contextualized communication and performance based on organizational competence in appropriate way (Bachman, 1990). In 1978, Munby highlighted the paramount importance of pragmatic knowledge and argued that in order to communicate effectively, a speaker has to have the ability to produce grammatical utterances of a language as well as the ability to use them appropriately. Swain (2005), in a similar vein, contends that language learners have to learn the social and pragmatic rules of the target-language and stresses that language teachers need to take pragmatics into consideration while teaching learners to communicate in another language.

Pragmatic failure, is described (Hudson et al., 1992) as circumstances in which the speaker's utterance is misunderstood by the hearer and this misunderstanding is caused, among other things, by inapt realization of speech acts. Speech acts are defined as utterances that are performed with intentions to have effects on listeners (Crystal, 1997). Refusals, requests, and apologies are among speech acts that necessitate the exploitation of various strategies depending on the culture (Beebe & Takahashi, 1989; Beebe et al., 1990), given the differences between perceptions of speakers of different languages due to different concepts such as power relations, social distance, and degree of imposition (Hudson et al., 1992).

Relative power is concerned with the degree of imposing the intention of the speaker on the hearer. Social distance, on the other hand, is defined as a function of the level of acquaintance between the two parties of communication. Thus, communicating with an unknown person would entail distance. Furthermore, the degree of imposition refers to the level of imposition of action on the hearer. Hence, situations demanding verbs such as asking for help would involve imposition. Due to the fact that speakers from various cultures perceive these variables differently (Hudson et al., 1992), language learners might fail to appropriately produce speech acts, which in turn would lead to pragmatic failure in communication. Given, the important role pragmatic competence plays in one's knowledge of how to communicate (Bachman, 1990; Bachman & Palmer, 2010), language instructors and researchers should seek ways to facilitate the acquisition of the above-mentioned variables, speech act and pragmatic competence. Therefore, delving into the ways

classroom practices can be manipulated so as to enhance language learners' organizational and pragmatic competence is worthy and justified. One of the classroom practices which have always been part of language teaching courses is classroom assessment. Nowadays, assessment is considered as a facilitator of learning and is deemed as a process rather than a product. Peer and self-assessment, which are two alternatives to exclusive assessment by the teacher, allow the involvement of learners in the assessment process and can lead to more successful teaching and learning process (e.g., Abolfazli Khonbi & Sadeghi, 2015; Chen, 2006; Hariri Asl & Marandi, 2017). Therefore, the notions and elements of assessment and specified criteria should be introduced to learners, and they should be trained how to evaluate their own and others' contributions via self and peer assessment (Wiliam & Thomson, 2007). Regarding assessment techniques as viable pedagogical practices, the present study sets out to delineate how making adjustments in assessment techniques can affect Iranian English as a foreign language (EFL) learners' organizational and pragmatic competence reflected in their oral productions requiring the use of speech acts.

Literature Review

Language Competence

Many applied linguists have contributed to defining the knowledge of using language; however, there is no agreement on the exact term to employ for such competence. To some (e.g., Canale, 1983, 1984; Canale & Swain, 1980; Widdowson, 1983), the term *communicative competence* could best describe language users' knowledge and skills to use the language, while to some others (e.g., Bachman, 1990; Bachman & Palmer, 1996, 2010), *communicative language ability* or *language competence* can best define both language proficiency and communicative competence. Recent theoretical and empirical enquiries into communicative competence mostly draw on the model promulgated by Canale and Swain (1980) and that of Bachman and Palmer (1996, 2010). Canale and Swain (1980) proposed three main components of grammatical, sociolinguistic and strategic competence for communicative competence, defining grammatical competence as being concerned with learning the linguistic aspects including vocabulary knowledge and morphological, syntactic, semantic, phonetic, as well as the knowledge of features of handwriting. To Canale and Swain, the sociolinguistic competence comprises of knowledge of rules and conventions which dominate the appropriate comprehension and language use in different contexts. In Canale and Swain's model, strategic competence entails knowledge of verbal and non-verbal communication strategies exploited to compensate for breakdowns in communication resulting from inadequate knowledge in components of communicative competence.

Bachman (1990) and Bachman and Palmer (1996, 2010) put forward a more inclusive model of communicative competence, drawing on the findings of prior empirical research, and theoretical postulations. The model of communicative competence suggested in 1990 by Bachman was later on modified by Bachman and Palmer in the mid 1990s and 2010. They argue that language ability comprises two broad areas which are language knowledge and strategic competence. Language knowledge consists of two main parallel components i.e. organizational knowledge and pragmatic knowledge which complement each other in effective use of language

for communication purposes. Organizational knowledge includes individuals' ability to control formal language structures, i.e., of textual and grammatical knowledge. Grammatical knowledge further encompasses knowledge of vocabulary, morphology, syntax, phonology, and physical features of handwriting. Textual knowledge, on the other hand, includes the knowledge of meaningful relationships among sentences (cohesion) and knowledge of developing different genres of language as well as conversational organization (conventions for opening, continuing and closing conversations).

The second component of language knowledge, pragmatic knowledge, refers to the knowledge language users need to enable them to generate and comprehend discourse by making connections among utterances and texts to their meanings as well as the intentions of language users and the relevant features of the language use context and setting (Bachman & Palmer, 2010). Pragmatic knowledge is, therefore, needed to generate and interpret discourse, comprising of two areas of knowledge namely, functional knowledge (i.e. knowledge of pragmatic conventions needed to produce proper language functions and interpret the illocutionary power of utterances/sentences) and sociolinguistic knowledge (i.e. knowledge of sociolinguistic conventions required to create and interpret contextually appropriate language utterances/sentences). The second component of language ability, strategic knowledge, encompasses a set of metacognitive components enabling language user to set goals, evaluate communicative sources, and plan. The study at hand focuses on the language knowledge of Bachman and Palmers' (1996, 2010) model, concentrating on pragmatic knowledge mirrored in the use of contextually proper utterances with regard to power, distance, and imposition, as well as organizational knowledge reflected in the demonstration of the ability to use error free use of vocabulary, grammar, and phonological presentations, and generating coherent utterances.

Assessment as a Learning Tool

One of the activities instructors are often asked to take on is assessment which encompasses "any process that provides information about the thinking, achievement or progress of students" (Crooks, 2001, p. 1). Though traditionally employed merely for judgmental purposes, assessment is nowadays viewed as a potential tool to aid and facilitate learning (e.g., Darling-Hammond et al., 2020; Lee, 2007; Taras, 2008). Recent approaches to assessing learning product and process have stipulated learners as their focus, highlighting and emboldening the involvement of learners in the assessment process (Darling-Hammond et al., 2020; Sluijsmans et al., 2003).

As opposed to traditional, mainly teacher-centered assessment techniques, more recent techniques are learner-oriented suggesting alternative methods of assessing learning particularly by learners themselves. Such learner-centered approaches to assessment are reported to offer more accurate evaluation of learners' ability (e.g., Abedi, 2010; Coombe et al., 2007), advocating a key role for learners in decisions pertinent to assessment and the undertaking of assessment procedures. Self-assessment and peer-assessment in which learners are asked to monitor their own or a peer's performance on a language learning task (Richards & Schmidt, 2010; Topping, 2009) are pigeonholed branded as such pioneering learner-oriented assessment approaches. Despite the recent popularity of alternative approaches to assessment which encourage the engagement of learners in assessment (Soleimani &

Rahmanian, 2014) their influence on learning is still under-researched. Assessment as a learning tool is yet a relatively under-probed area in EFL contests, and scant research has been documented in the literature scrutinizing how alternative assessment techniques can assist the acquisition of another language. To date, most studies addressing alternative innovative assessment techniques mostly target to rationalize and defend their legitimacy as assessment tools (De Saint-Léger, 2009). Among the studies carried out to examine the benefits of peer and self-assessment techniques are those intended to assess the impact of peer and self-assessment on goal-orientation (Zarei & Yousefi, 2015), reading motivation (Rahmany et al., 2013; Wolters, 1999), reading comprehension (Shams & Tavakoli, 2014), writing skill (Birjandi & HadidiTamjid, 2010; Birjandi & Siyyari, 2010; Soleimani & Rahmanian, 2014; Williams, 2012), creating social presence in conjunction with CALL (Hariri Asl & Marandi, 2017), the relationship between EFL students self-perceived communication competence and their task-free and task-based self-assessment of speaking (Abolfazli Khonbi & Sadeghi, 2015), and learner autonomy (Ashraf & Mahdinezhad, 2015). As the above account of studies on the potential role of alternative assessment techniques in learning suggests, despite the widely acknowledged standing of communication as realized by educationalists in the twenty-first century (Purpura, 2017; Richards, 2008), to date, to the best of our knowledge, few studies have examined the potential effect of alternative assessment on communicative competence as their objective. In 2014, Pakzadian and Tajeddin investigated the impacts of instruction that was based on self-assessment on 30 language learners' acquisition of three types of speech acts including suggestion, complaint, and request. The results of their study revealed that although the self-assessment group outperformed the control group on most of the items in a written discourse completion test, their overall score did not have a significant difference with the comparison group. Therefore, they introduced self-assessment as an ineffective task for acquiring pragmatic knowledge and meta pragmatic awareness. They attributed the results to the point that self-assessment process in their study was more concerned with problems in answering the discourse test rather than teaching students how to communicate appropriately and effectively using the pragmatic features of language. Therefore, it seems that more studies are needed to provide a clear picture of the role of learner-centered assessment techniques in acquiring pragmatic knowledge.

The Present Study

Given the scarcity of research on the benefits that alternative assessment techniques can offer to facilitate the acquisition of L2 in general and the gap in the literature concomitant with the role of peer and self-assessment techniques in promoting language knowledge as defined by Bachman and Palmer (1996, 2010), the present study aimed at investigating the differential effect of various assessment techniques on learners' pragmatic and organizational competencies. In so doing, the ensued null hypotheses were formed:

1) Different assessment techniques (i.e., self, peer, and teacher-assessment) do not have any significant effect on participants' pragmatic competence and learners who benefited from self-assessment, peer-assessment, and teacher-assessment have similar levels of pragmatic competence all other conditions maintained alike.

2) Different assessment techniques (i.e. self, peer, and teacher-assessment)

do not have any significant effect on participants' organizational competence and learners who benefited from self-assessment, peer-assessment, and teacher-assessment have similar levels of organizational competence all other conditions maintained alike.

To further test the null hypotheses, the following research questions were posed:

Do various assessment techniques (i.e. self, peer, and teacher-assessment) have any significant effect on participants' pragmatic competence?

Do various assessment techniques (i.e. self, peer, and teacher-assessment) have any significant effect on participants' organizational competence?

Method

Due to non random sampling, the study enjoyed a quasi experimental design including pre-and post-tests. The dependent variables were pragmatic and organizational knowledge and the independent variables were self, peer, and teacher-assessment techniques.

Participants

A total number of 98 female Iranian learners of English studying at pre-intermediate levels in an Iranian language institute with an age range of 15 to 29 (M = 17.2, SD = 8.32) participated in this study. The participants were chosen from among 157 pre-intermediate learners who had taken Cambridge's Preliminary English Test (PET) (M = 53.46, SD = 6.92) and scored one standard deviation from mean. In other words, the administration of PET abetted the researchers to ensure the homogeneity of the participants in terms of general linguistic knowledge. The first 32 learners were randomly assigned to the self-assessment group (SA), in which participants were instructed to evaluate their own performance and learning, and 33 to the peer-assessment group (PA), where learners were taught how to assess their peers. Finally, the last group which included 33 learners was the teacher-assessment group (TA), in which the assessment of the learners' performance in the class was conducted merely by the teacher.

Instruments

Speech acts involving the element of power relation can be more effectively assessed through role-play (Grabowski, 2013; Kasper & Rose, 2001, 2002; Purpura, 2017). Therefore, to assess pragmatic knowledge, Discourse Role Play Talks (DRPTs) were employed in this study. Employed mostly as a tool for pedagogical ends across a wide range of subjects, role plays are also used for assessment purposes in the evaluation of speaking and pragmatic competence (Kasper & Youn, 2018). Role plays provide a context to examine how language learners generate and comprehend utterances while in interaction, speculating a specific context for learners and thus ensuring that not only does the speaking test enjoy authenticity but also it is not afflicted by the under-representation of the construct under examination. In DRPTs, students are required to read the explanation of a situation and to play a role based on the same situation with either a teacher or a peer. In this way, learners are mandated to come up with sentences commensurate with the situation described both in terms of structural and pragmatic considerations. For each scenario learners were given 8 to 12 minute to conduct the role play. They were given two minutes to prepare for the role play and think over

the situation and the language and function they demanded prior to role plays. However, they could not ask for the researchers' assistance while engaged in the role plays. Nor were they given any cues as whether their utterances were correct or appropriate.

Three speech act situations were postulated as role plays, in which participants were required to communicate with either a peer or the teacher as interlocutor. The role play test comprised of eight different scenarios adopted from Enochs and Yoshitake-Strain (1999), each including all three speech acts namely, a request, a refusal, and an apology with different degrees of power, imposition and distance in various situations. Request, refusal and apology were selected as the main functions because they happen often in everyday speech acts and they were among the functions that had to be taught to the participants of the present study. Furthermore, due to their significant role in the pragmatic knowledge of EFL learners, these three speech acts have commonly been researched (e.g., Birjandi & Derakhshan, 2014; Felix-Brasdefer, 2006, 2008), although the focuses of these studies have been on different aspects. For example, Birjandi and Derakhshan investigated the effect of video-driven prompts on promoting students' awareness of the Pragmatic comprehension of apology, request and refusal. Another important feature of these acts is that they can be expressed in several different ways. In this regard, Purpura (2017) refers to the well-known example of a person requesting for salt in a restaurant and mentions that the request can be friendly or unfriendly, patient or humorous, and demanding or sarcastic. Therefore, depending on the degrees of power, distance and imposition, request, refusal and apology can be expressed in several different ways.

Written descriptions of the role plays including what was supposed to be done by the interlocutor was presented to the participants in both Persian and English to elicit the use of speech acts in the role play. Learners' performances in the role plays were video-taped for the sake of analysis in terms of pragmatic and organizational knowledge.

Assessing Pragmatic Knowledge

To gauge the appropriateness of the generated utterances, having watched the recorded role plays, two experienced English language teachers assessed each individual's performance with regard to each speech act in each role play. They used a 6 point Likert scale, with 1 indicating the least appropriate and six signaling the most appropriate performance. The scores gained for all speech acts comprised the participant's score in that specific scenario, and the scores gained in all eight scenarios were tallied and added up to create the participant's pragmatic competence score in an administration of the DRPT. Interrater reliability was established through calculating correlation coefficient of the scores awarded by the two raters. This coefficient was found to be .79 in the pretest. The two raters resolved the differences through discussion, listening to video-taped role plays again, which yielded to the achievement of the inter-rater reliability coefficient estimated in the posttest was .86.

Assessing Organizational Knowledge

In order to delineate the organizational knowledge of participants, four statements were employed to form a 6 point Likert-scale to rate each participant's performance in each role play. These statements were to cover the two basic areas of organizational knowledge in Bachman and Palmer's model, grammatical and textual

knowledge. Therefore, the first sentence stated that: "the utterances were error free in terms of grammar". Adopted from the inventory of grammatical areas identified in the Cambridge English PET handbook for teachers, the grammatical structures expected to be appropriately used for pre-intermediate participants were as follows: Modals: can (for example, the learners had to be able to use can, based on the context, in three different ways: can = ability; can = request; can = permission). Another example of modals with three possible meanings was could as ability, could as possibility and could as polite request. The other grammatical areas included: would (polite requests) will (offer) shall (suggestion; offer) should (advice) may (possibility) might (possibility) have (got) to (obligation) ought to (obligation) must (obligation) mustn't (prohibition) need (necessity) needn't (lack of necessity) used to + infinitive (past habits); Tenses : present simple, present continuous, present perfect simple, past simple, past perfect simple, future with going to, future with present continuous and present simple, future with will and shall; Verb forms: Affirmative, interrogative, negative imperatives, infinitives, gerunds after verbs and prepositions, gerunds as subjects and objects, passive forms, causative have/get, So/nor with auxiliaries; Conditional sentences Type 0, 1, and 2; Simple reported speech, indirect and embedded questions. Therefore, erroneous statements beyond these grammatical structures (e.g., an error in the formation of conditional type 3) were tolerated and not regarded as an error for the participants of the present study since such structures are deemed to be beyond their organizational knowledge. The second sentence was: "The utterances were error free in terms of vocabulary use". The third and the fourth sentences consecutively said that: "There were no mistakes pertinent to pronunciation and intonation" and "Utterances were coherently related and proper cohesive devises were used". The scale was to be answered by two researchers individually, with 6 indicating "strong agreement" and 1 signaling "strong disagreement". Interrater reliability was measured through calculating correlation coefficient of the two raters' scores. In the pretest, this coefficient was calculated as .72, which was increased to .81 in the posttest before which the two raters had resolved the differences having listened to taped role plays again and discussed differences of view point.

Procedure

In order to be able to gauge the difference of the learners' performance in pre and posttest, prior to the treatment, the researchers administered the Discourse Role Play Talks (DRPTs) to all learners and asked them to act out the situations in the cards with peers. The pair for the role plays were selected and assigned based on random selection and learners were not permitted to choose their own peers for the role plays. The performance of the participants in the role plays both in the pretest and in the posttest were recorded to allow the analysis of the sentences in terms of pragmatic competence and organizational knowledge. Having watched the recorded performance of the participants, two of the researchers rated the participants using the pragmatic and organizational knowledge scales explained in the previous sections. The administration of the DPRT pre-test was followed by the treatment which lasted six 45-minute sessions.

In each of these sessions, the participants in the self-assessment (SA) group were presented with the two scales employed as the instruments in this study to assess pragmatic and organizational knowledge. They were briefed about the scoring and the criteria based on which they would be assessed in terms of

pragmatic as well as organizational competences. In so doing, the statements in both scales were explained to them and clarified using examples. They were told what each item meant and on what ground the performance was to be rated. Next, on the same session, participants were presented with the role play scenarios each of which they had to act out with a peer every session. With participants' consent, their performance was videotaped, so that participants would be able to review, reflect on, and analyze and assess their own performance. Having acted out the scenarios in the role plays, they were asked to answer the scales and evaluate their own performance. They did the assessment independent from that of their peers in the role play, while watching the pertinent video. They were then free to delete the video file if they desired. They were instructed to collect these answered scales in a portfolio which had to be submitted to researchers prior to the posttest to provide evidence for the researchers that the self-assessment had been carried out.

The peer assessment (PA) group participants, in the same vein, received the two scales and the same role play scenarios in which they were asked to act out with a peer each session while being video-taped. Similar to the learners in self-assessment group, the peer assessment group participants were not allowed to choose the partners for the role plays. They, however, unlike their SA counterparts, were asked to assess the performance of their peer, and not that of their own, in each role play concerning pragmatic and organizational competence and accumulate these evaluations in a portfolio to be handed in to the researchers at the end of the treatment to prove their peer assessment activities.

Unlike SA and PA participants, the teacher assessment (TA) group learners did not get the chance to get involved in the assessment process. They did not receive the assessment scales either. They were merely assigned to pairs to act out the scenarios each session. Having done their role pays, they were assessed by the teacher in terms of pragmatic and organizational competence and were later informed about their total organizational and pragmatic knowledge scores that had been decided by the teacher.

The treatment was followed by another PDRT as the post test. The role plays of the post-test were different from the ones in the pre-test; however, the same procedure of scoring and the same scales used by the same raters in the pre-test were employed to gauge the participants' pragmatic and organizational knowledge.

Results

To demarcate the impact of assessment type on participants' pragmatic and organizational competence reflected in their performance in role-plays focused on three speech acts, the researchers scrutinized data gleaned from the pragmatic knowledge and organizational knowledge scales. First, a normality test was run to make sure the normality condition was met. The result of a Kolmogorov-Smirnov and Shapiro-Wilk revealed no violation of normality in pre- and post-tests ($p > .05$). Then, data accumulated from each scale were separately analyzed. Gains from pre-test to post-tests were calculated through subtracting pretest scores from post-test scores and were subjected to one-way Analysis of Variance (ANOVA).

Pragmatic Competence

The first research question in the current study was intended to scrutinize whether assessment techniques had any significant effect on participants' pragmatic competence as mirrored in the appropriateness of the use of three speech acts in the

Discourse Role Play Talks (DRPTs) test. The first null hypothesis in the present study stated that different assessment techniques do not have differential impact on participants' pragmatic competence. To examine the null hypothesis and to find out whether pragmatic competence could be a function of assessment techniques, first, the researchers employed descriptive statistics to calculate means and standard deviations of the groups in both pre-test and post-test, the results of which are illustrated in Table 1.

Table 1
Descriptive Statistics for Pragmatic Competence

		Pre-test		Post-test		Gain	
Groups	N	Mean	SD	Mean	SD	Mean	SD
SA	32	80.46	9.61	115.15	17.92	34.68	9.29
PA	33	70.93	7.35	90.69	13.43	19.75	6.68
TA	33	76.60	8.62	82.54	10.49	5.93	5.05
Total	98	75.95	9.34	95.93	19.76	19.97	13.75

As demonstrated in Table 1, when gain scores of the pragmatic test were calculated, it was found that learners who had received the opportunity to play a part in assessment procedure (i.e. SA [M = 34.68, SD = 9.29], and PA [M = 19.75, SD = 6.68] groups' participants) obtained higher gain scores compared to the TA group participants (M = 5.93, SD = 5.05) who had not been granted the opportunity to assess their own performance or that of a peer throughout the experiment.

Table 1 also reveals that the learners who were given the responsibility to critically review and evaluate their own performance (i.e. SA [M = 34.68, SD = 9.29]) gained the highest gain scores in the pragmatic test. Next, an ANOVA was run on the gain scores to find out whether the observed differences were statistically significant. Shown in Table 2 are the results of the one-way ANOVA on pragmatic competence gain scores.

Table 2
One-way ANOVA, Pragmatic Competence

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Between groups	13429.14	2	6714.57	129.36	.00
Within groups	4930.81	95	51.90		
Total	18359.95	97			

As shown in Table 2, significant differences were observed among the gain scores of participants in the SA (M = 34.68, SD = 9.29), PA (M = 19.75, SD = 6.68), and TA (M = 6.68, SD = 5.05) groups, in the pragmatic competence test, $F(2, 97) = 129.36$, $p < .05$.Then, to find the differences, a post hoc test was run. The results of a post hoc Scheffe's test, illustrated in Table 3, further elucidated where the significant differences were located.

Table 3
Scheffe's Test: Pragmatic Competence

Group	PA	TA
SA	14.92*	28.74*
PA		13.81*

*. The mean difference is significant at the 0.05 level.

As Table 3 suggests, both alternative assessment techniques (i.e. the self-assessment and the pear-assessment) were found to yield significant increase in participants' pragmatic competence since the results of the post hoc test indicated that gain scores of the SA and PA groups were significantly higher when compared to that of the group the members of which were assessed merely by the teacher (i.e. TA group). The first null hypothesis was, therefore, rejected. Furthermore, the post hoc test crystallized a significant mean difference between the gain scores of the SA and PA participants when posttest pragmatic scores were checked against pre-test ones, suggesting that taking control of the process of the assessment could enhance participants' pragmatic knowledge as reflected in the proper use of speech acts in role play scenarios more than peer-assessment, and thus, introducing self-assessment as the most effective assessment technique for promoting pragmatic competence. Moreover, the learners who evaluated the performance of their fellow classmates in the role plays, the SA group, also significantly scored higher than the TA group, which indicated that having a say in the assessment procedure of the evaluation of pragmatic competence, even if not the assessment of one's own performance, can have desired impact on pragmatic knowledge.

Organizational Competence

The second research question addressed the effect of assessment technique on participants' organizational knowledge which was operationalized as the accurate use of vocabulary and structures, as well as proper pronunciation of the well-linked and cohesive utterances. To find the answer to the second research question, and to see if the second null hypothesis stating that different assessment techniques do not have differential impact on participants' organizational competence could be rejected, the researchers first calculated means and standard deviations of the groups in the two administrations of the organizational competence test. Table 4 depicts descriptive statistics pertaining to organizational competence tests.

Table 4
Descriptive Statistics for Organizational Competence

		Pre-test		Post-test		Gain	
Groups	N	Mean	SD	Mean	SD	Mean	SD
SA	32	14.15	2.84	21.28	2.60	7.12	3.46
PA	33	14.66	2.64	20.06	1.78	5.39	2.66
TA	33	14.00	2.43	16.33	2.73	2.33	1.49
Total	98	14.27	2.63	19.20	3.18	4.92	3.29

As Table 4 indicates, regarding organizational competence, alternative assessment techniques groups (i.e., SA [M = 7.12, SD = 3.46], and PA [M = 5.39, SD = 2.66] groups) were found to have improved more from the pre-test to the post-test when compared to the learners in the TA group (M = 2.33, SD = 1.49) that were evaluated by their teacher. Further, similar to the pattern observed in pragmatic competence test, participants who had evaluated their own organizational knowledge in role plays obtained the highest gain scores. The researchers, then, ran an ANOVA on the organizational competence gain scores to establish the significance of the observed differences. Displayed in Table 5 are the results of the one-way ANOVA on organizational competence gain scores.

Table 5
One-way ANOVA, Organizational Competence

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Between groups	383.78	2	191.89	27.18	.00
Within groups	670.71	95	7.06		
Total	1054.50	97			

Table 5 indicates significant differences among the gain scores of participants in the SA (M = 7.12, SD = 3.46), PA (M = 5.39, SD = 2.66), and TA (M = 2.33, SD = 1.49) groups, in the organizational competence test, $F(2, 97) = 27.18$, $p < .05$. The second null hypothesis was, therefore, rejected. A post hoc Scheffe's test was then employed to locate the significant differences. Table 6 illustrates the results of the post hoc test on organizational competence gain scores.

Table 6
Scheffe's test: Organizational Competence

Group	PA	TA
SA	1.73*	4.79*
PA		3.06*

*. The mean difference is significant at the 0.05 level.

As shown it in Table 6, alternative assessment techniques encouraging learners' involvement in the testing procedure significantly improved participants' organizational competence as the SA and PA groups gain scores were found to be significantly higher than the scores of the TA group. Additionally, the post hoc test revealed a significant mean difference between the gain scores of the SA and PA participants in the organizational competence, which rendered self-assessment more effective in promoting organizational knowledge as measured by the precision of vocabulary and structure use, correctness of pronunciation, and cohesion of utterances.

Discussion and Conclusion

The main objective of this study was to find out whether self and peer-assessment techniques could be employed as instructional tools to improve pragmatic and organizational competences as the two main components of language knowledge as postulated by Bachman and Palmer (1996). The study set out to ascertain whether getting involved in the evaluation processes (e.g., of one's own performance or a peer's) could significantly improve pragmatic knowledge, reflected in the correct and appropriate use of three speech acts in role plays. Another objective of the study was to scrutinize the effect of peer and self-assessment on organizational knowledge, mirrored in phonologically, semantically, and syntactically accurate and cohesive utterance.

The results of the ANOVA tests revealed that the learner-centered assessment techniques (i.e. self and peer-assessment) could significantly improve both pragmatic and organizational knowledge. Evaluating one's own oral performance was proved to be more effective than assessing the performance of a peer as the results proved that the learners in the self-assessment group had outperformed those in the peer assessment group in both pragmatic and organizational competence scales.

The results pointed out that engaging learners in the assessment procedure

can desirably affect their knowledge of pragmatics which plays a paramount role in communication. Put differently, having been familiarized with the procedures and criteria of assessment of pragmatic knowledge, and through assessing themselves and their peers, the learners performed better in the pragmatic competence posttest. Organizational knowledge was also found to be promoted as the result of providing the learners by the opportunity to take part in assessment decisions. This result crystallizes the pedagogical value of learner-centered assessment techniques.

The results can be explained in the light of the awareness that self and peer assessment groups gained. They received information about the scales and became acquainted with the criteria based on which their performance was to be gauged. It seems that by being given the opportunity to closely study the assessment criteria, they were able to set specific goals for themselves which enabled them to achieve more (Schunk, 1989). Learners can comprehend the process of reaching objectives when they can evaluate their work (Kitsantas et al., 2004; Shih et al., 2010). Therefore, presenting evaluation criteria assists learners in gaining a thorough understanding of outcomes and expectations as well as an awareness of the ways to internalize the steps necessary to meet the goals (McMillan & Hearn, 2008). In the same vein, participating in self and peer assessment groups could have assisted the learners of the present study to become aware of their strengths and weaknesses, paving the way for creating autonomous learners. It seems that assessments other than teacher assessment are promising moves towards integrating teaching/learning in line with the approach advocated by Turner and Purpura, (2016) as learning oriented assessment (LOA) approach. Furthermore, the results confirm the sociocultural approaches to language learning and assessment in which, in accordance with Vygotskian dynamic assessment concepts (Poehner et al., 2017), teaching is an inherent part of all assessments demanding a close cooperation between the assessor and learners.

Self-assessment and peer assessment, as important components of formative assessment, arm learners with knowledge of the expected and desired performance as well as the criteria on which evaluation is grounded. Rather than merely asking learners to give a score or grade themselves, such alternative assessment techniques are able to encourage and stir the making of qualitative judgments about one's own performance (Andrade, 2010). Adopting such an active role in the assessment process enables individuals to gain deeper understanding of their current level at a particular aspect of the targeted skill and elevate it, urging them to bridge the gap between a current performance (as reflected in the result of the assessment) and the desired level. Also, through elucidating the purpose of the performance, such involvement with assessment can prompt learners to monitor their learning process and make adjustments when required (Panadero et al., 2017). Thus, they can foster self-regulated learning (Andrade, 2010) which can in turn promote learning itself (Zimmerman, 2011). As Goldman and Pellegrino (2015) put it, when learners collaborate in peer assessment, they become aware of each other's thinking and find the opportunity to share their perspectives and strategies. Consequently, they may provoke and broaden each other's thinking and understanding. They add that self-assessment can contribute to students' monitoring and controlling their own learning by enabling them to identify their lacks and needs and to adjust their learning activities accordingly. It seems that learners who are given only the teachers' assessment results and are expected to follow the feedbacks with no perception of their aims, will not acquire the skill to monitor and self-adjust

their performance and, consequently, cannot promote their skills as much (Panadero et al., 2017).

The impact of assessment techniques on various components of language knowledge, particularly pragmatic competence, had not been examined in previous studies. Thus, while adding to the literature through filling this gap, the results corroborate the ones gained in the line of studies reporting the superiority of alternative assessment over traditional teacher-centered methods of evaluation in language learning (e.g., Harris, 1997; Meisels et al., 2003).

These results are also in line with some studies such as Butler and Lee (2010), Pitsoulakis and Bailey (2016), and Ross (2005), although they scrutinized areas other than pragmatic knowledge. For example, Butler and Lee (2010) investigated the effectiveness of self-assessment on a rather large number of young EFL learners in South Korea. The participants practiced self-assessment in their English language classes for one semester and at the end both their self-assessment skills and language skills improved. Ross (2005) did a longitudinal study and compared the effect of summative and formative assessments on the language proficiency of EAP students and realized that formative assessment practices including peer and self-assessments yield better results. Furthermore, the results are partly in line with those of Hariri Asl and Marandi, (2017) who concluded that in online learning communities that attempt to create social presence, peer assessment can yield more fruitful results followed by self and teacher assessment. The same agreement can be found between the results of this study and Birjandi and Siyyari's (2010) study. They investigated the effect of peer and self-assessment on the learners' paragraph writing performance and concluded that peer and self-assessment are more fruitful than teacher assessment, although in their study peer assessment came first.

The results are not in line with the ones gained by Pakzadian and Tajeddin (2014) who reported that, in spite of the better performance of the students in the self-assessment group in the written completion test, they did not outperform the control group overly. This discrepancy can be explained with regard to the differences in assessment tools. In the study at hand, role plays, which can be a direct measure of pragmatic competence, were applied, whereas Pakzadian and Tajeddin (2014) applied a written discourse completion test to assess pragmatic knowledge.

As one of the main objectives of any language learning program is being able to use it effectively to interact and communicate, pedagogical practices in language classrooms, including assessment methods, should be designed in a way as to trigger and encourage critical evaluation of performance which has proved to significantly impact two main components of the ability to communicate. All in all, the results of the present study imply that self and peer assessment ought to be among pedagogical practices in language classroom, particularly when the speaking skill is the target. Given the importance of pragmatic considerations (Kasper & Rose, 2001) and regarding the fact that assessment of one's own performance (or that of a peer) provides a deeper understanding of the component under study, self-assessment and peer-assessment interventions can have a desired impact on students' self-regulated learning of the pragmatics, leading to the empowerment of learners in acquiring pragmatic component of communicative competence in interactions.

The present study contributed to the still limited body of research into the

pedagogical convenience of alternative assessment techniques as it included the examination of variations in assessment techniques in the promotion of two main components of language knowledge and introduced self and peer-assessment techniques as vital learning tool assisting instructors to elevate both organizational and pragmatic competences. The results of the current study are expected to motivate language teachers to encourage self-evaluation and self-assessment techniques and familiarize learners with assessment criteria and steps and involve them in assessment procedures. It is worth mentioning that although the results indicated that self and peer assessment yield more fruitful results than teacher assessment, concerns about the cost and time of implementation have to be considered. When put into practice, there are issues pertinent to successful implementation of self and peer assessment. Students need clear guidelines and close supervisions and providing these guidelines and supervisions can be time consuming specially in classes with large number of students. Furthermore, some students might resist the alternative assessments, because they are used to traditional assessments performed by their teachers. Teachers' knowledge of assessment methods is another concern. Tajeddin et al. (2018), in a study about novice and experienced teachers' assessment literacy concluded that these teachers "have both shared and divergent speaking assessment literacy" (p. 57) and that the past education as well as their years of teaching experiences can be the determining factors in their knowledge and awareness. Therefore, the training of teachers and methods of raising their awareness with regard to effective ways of implementing self or/ and peer assessment techniques should be the focus of future studies and a main concern for educators. Studies are needed to compare and delve into different ways in which assessment done by learners can be introduced to the learning context and the ways learners react and respond to them.

The present study suffered a few limitations and delimitations that mainly pertained to the participants. As mentioned earlier, the participants in the study were ninety- eight female English learners studying at pre-intermediate levels with an age range of 15 to 29. The fact that participants were selected from merely one proficiency level can be regarded as a limitation of the study. Assessing the effect of treatment on pragmatic knowledge across various proficiency levels would have culminated in a more thorough understanding of the phenomenon under study. Moreover, the inclusion of participants from both genders might have illuminated the possible moderating role of gender, thus providing a bigger picture. In addition, the limited number of the sessions in the treatment, mainly due to working on the oral performance of the subjects, can be regarded as another limitation of the present study.

Studies employing various methods of data gathering are needed to confirm the obtained results in written productions of language learners. Learners' generated texts can be analyzed for pragmatic and organizational appropriateness and accuracy to shed light on the effectiveness of self- and peer-assessment in promoting language knowledge. The effect of assessment techniques on strategic competence is also to be scrutinized in further studies. Finally, further studies are needed to ensure the validity and reliability of alternative assessments in different contexts.

References

Abedi, J. (2010). Adapting performance assessment for English language learners. In L. Darling-Hammond & F. Adamson(Eds.), *Beyond the bubble test: How performance assessments support 21st century learning*, (pp. 185-206). John Wiley & Sons.

Abolfazli Khonbi, Z., & Sadeghi, K. (2015). Self-, peer-, and teacher-assessment: An investigation into Iranian EFL students' attitudes. *Studies in Second Language Learning and Teaching*, *3*(1), 87-107. https://doi.org/10.14746/ssllt.2013.3.1.5

Andrade, H. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. J. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90-105). Routledge.

Ashraf, H., & Mahdinezhad, M. (2015). The role of peer-assessment versus self-assessment in promoting autonomy in language use: A case of EFL learners. *Journal of Language Testing, 5*(2), 104-113. https://www.ijlt.ir/journal/images/PDF/437-2015-5-2.pdf

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford University Press.

Beebe, L., & Takahashi, T. (1989). Sociolinguistic variation in face-threatening speech acts: Chastisement and disagreement. In M. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language variation* (pp. 199-218). Plenum Press.

Beebe, L., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. C. Scarcella, E. Andersen, & S. D. Krashen (Eds.), *Developing communicative competence in a second language, (pp. 55-73)***.** Newbury House.

Birjandi, P., & Derakhshan, A. (2014). Pragmatic comprehension of apology, request and refusal: An investigation on the effect of consciousness-raising video-driven prompts. *Applied Research on English Language, 3*(1), 67-86. https//doi.org/10.22108/ARE.2014.15479

Birjandi, P., & HadidiTamjid, N. (2010). The role of self-assessment in promoting Iranian EFL learners' motivation. *English Language Teaching, 3*(3), 211-220. https//doi.org/10.5539/elt.v3n3p211

Birjandi, P., & Siyyari, M. (2010). Self-assessment and peer-assessment: A comparative study of their effect on writing performance and rating accuracy. *Iranian Journal of Applied Linguistics (IJAL)*, *13*, 23-45. https://ijal.khu.ac.ir/article-1-44-en.html

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing, 27*(1), 5-31. https://doi.org/10.1177/0265532209346370

Canale, M. (1983). A communicative approach to language proficiency assessment in a minority setting. In C. Rivera (Ed.), *Communicative competence approaches to language proficiency assessment: Research and application*, (pp. 107-122). Multilingual Matters.

Canale, M. (1984). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication,* (pp. 2-27). Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1-47. https://doi.org/10.1093/applin/i.1.1

Chen, Y. M. (2006). Peer- and self-assessment for English oral performance: A study of reliability and learning benefits. *English Teaching and Learning*, *30*(4), 1-22.

https://doi.org/10.6330/ETL.2006.30.4.01

Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners*. University of Michigan. https://doi.org/10.3998/mpub.170430

Crooks, T. (2001, September 13-15). *The validity of formative assessment* [Paper presentation]. The British Educational Research Association Annual Conference, University of Leeds. https://www.leeds.ac.uk/educol/documents/00001862.htm.

Crystal, D. (1997). *A dictionary of linguistics and phonetics* (4th Ed.). Blackwell.

Darling-Hammond, L., Flook, L., Cook-Harvey, Ch., Barron., B., & Osher, D. (2020) Implications for educational practice of the science of learning and development, *Applied Developmental Science, 24*(2), 97-140. https://doi.org/10.1080/10888691.2018.1537791

De Saint-Léger, D. (2009). Self-assessment of speaking skill and participation in foreign language class. *Foreign Language Annals, 42*, 158-178. https://doi.org/10.1111/j.1944-9720.2009.01013.x

Enochs, K., & Yoshitake-Strain, S. (1999). Evaluating six measures of EFL learners' pragmatic competence. *Japan Association for Language Teaching Journal (JALT)*, *21*(1), 29-50. https://doi.org/10.1111/j.1944-9720.2009.01013.x

Felix-Brasdefer, J. C. (2006). Linguistic politeness in Mexico: Refusal strategies among male speakers of Mexican Spanish. *Journal of Pragmatics, 38*(12), 2158-2187. https://doi.org/10.1016/j.pragma.2006.05.004

Felix-Brasdefer, J. C. (2008). Perceptions of refusals to invitations: Exploring the minds of foreign language learners. *Language Awareness, 17*(3), 195-211. https://doi.org/10.1080/09658410802146818

Goldman, S., & Pellegrino, J. (2015). Research on learning and instruction: Implications for curriculum, instruction, and assessment. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 33-41. https://doi.org/10.1177/2372732215601866

Grabowski, K. (2013). Investigating the construct validity of a role-play test designed to measure grammatical and pragmatic knowledge at multiple proficiency levels. In S. J. Ross & G. Kasper (Eds.), *Assessing second language pragmatics,* (pp. 149-171). Macmillan.

Hariri Asl, M. H., & Marandi, S. (2017). Peer-assessment and student-driven negotiation of meaning: Two ingredients for creating social presence in online EFL social contexts. *Issues in Language Teaching (ILT)*, *6*(1), 117-144. https://doi.org/10.22054/ILT.2017.8421

Harris, M. (1997). Self-assessment of language learning in formal settings. *English Language Teaching (ELT) Journal, 51*(1), 12-14. https://doi.org/10.1093/elt/51.1.12

Hudson, T., Detmer, E., & Brown, J. D. (1992). *A framework for testing cross-cultural pragmatics (Technical Report 2)*. University of Hawai'i, Second Language Teaching and Curriculum Center.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics,* (pp. 269-293). Penguin.

Ifantidou, E. (2013). Pragmatic competence and explicit instruction. *Journal of Pragmatics*, *59*, 93-116. https://doi.org/10.1016/j.pragma.2012.12.008

Ishihara, N., & Cohen, A. D. (2014). *Teaching and learning pragmatics: Where language and culture meet*. Routledge.

Kasper, G., & Rose, K. R. (2001). Pragmatics in language teaching. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 1-9). Cambridge University Press.

Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Blackwell.

Kasper, G., & Youn, S. (2018). Transforming instruction to activity: Role play in language assessment. *Applied Linguistics Review*, *9*(4), 589-616. https://doi.org/10.1515/applirev-2017-0020

Kitsantas, A., Robert, A., & Doster, J. (2004). Developing self-regulated learners: Goal setting, self-evaluation, and organizational signals during acquisition of procedural skills. *The Journal of Experimental Education*, *72*(4), 269-287. https://doi.org/10.3200/jexe.72.4.269-287

Lee, I. (2007). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *Canadian Modern Language Review*, *64*(1), 199-213. https://doi.org/10.3138/cmlr.64.1.199

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons, 87(1)*,40-49. https://files.eric.ed.gov/fulltext/EJ815370.pdf

Meisels, S., Atkins-Burnett, S., Xue, Y., & Bickel, D. D. (2003). Creating a system of accountability: The impact of instructional assessment on elementary children's achievement scores. *Educational Policy Analysis Archives. 11*(9), 1-18. https://doi.org/10.14507/epaa.v11n9.2003

Munby, J. (1978). *Communicative syllabus design*. Cambridge University Press.

Pakzadian, S., & Tajeddin, Z. (2014). Dynamic self-assessment: Is it an effective task for pragmatic development and meta pragmatic awareness? *Procedia - Social and Behavioral Sciences, 136*, 283-287. https://doi.org/10.1016/j.sbspro.2014.05.329

Panadero, E, Jonsson, A., & Botella, J. (2017) Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Pitsoulakis, D., & Bailey, A. L. (2016, April). *Extending learning progressions to self-assessment: Students finding "best fit" for next-steps learning and instruction* [Paper presentation]. The Annual Meeting of the American Educational Research Association, Annual Conference, Washington, DC.

Poehner, M. E., Davin, K. J., & Lantolf, J. P. (2017). Dynamic assessment. In E. Shohamy, L. G. Or, & S. May (Eds.), *Language teaching and assessment*, (pp. 243-256). Springer International Publishing.

Purpura, J. (2017). Assessing meaning. In E. Shohamy, L. G. Or, & S. May (Eds.), *Language teaching and assessment*, (pp. 33-62). Springer International Publishing.

Rahmany, R., Hassani, M. T., & Noroozi Larsari, V. (2013). The role of self-assessment in promoting Iranian EFL learners' motivation in reading English texts. *Life Science Journal*, *10*(7), 983-988. https://www.lifesciencesite.com

Richards, J. C. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson Longman.

Ross, S. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics, 26*(3), 317-342. https://doi.org/10.1093/applin/ami011

Schunk, D. H. (1989). Social-cognitive theory and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning and academic achievement: Theory, research and practice*, (pp. 83-110). Springer-Verlag.

Shams, N., & Tavakoli, M. (2014). The effect of peer, self, and traditional assessment on Iranian EFL learners' L2 reading comprehension. *Journal of Applied Linguistics and Language Research, 1*(1), 29-44.

Shih, K. P., Chen, H. C., Chang, C. Y., & Kao, T. C. (2010). The development and

implementation of scaffolding-based self-regulated learning system for e/m-learning. *Educational Technology and Society*, *13*(1), 80-93. https://www.learntechlib.org/p/75236/

Sluijsmans, D., Brand-Gruwel, S., van Merriënboer, J., & Bastiaens, T. (2003). The training of peer-assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation*, *29*(1), 23-42. https://doi.org/10.1016/s0191-491x(03)90003-4

Soleimani, H., & Rahmanian, M. (2014). Self-, peer-, and teacher-assessments in writing improvement: A study of complexity, accuracy, and fluency. *Journal of Research in Applied Linguistics*, *5*(2), 128-148. https://rals.scu.ac.ir/article_11016_994f2efb124ba6e4ccbddf3a1b517656.pdf

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning,* (pp. 471-483). Lawrence Erlbaum.

Sykes, I. M. (2013). Learner apologies in Spanish. *Technology in Interlanguage Pragmatics Research and Teaching*, *36*, 71-100. https://benjamins.com/catalog/lllt.36.05syk

Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. Oxford University Press.

Tajeddin, Z., Alemi, M., & Yasaie, H. (2018). Classroom assessment literacy for speaking: Exploring novice and experienced English language teachers' knowledge and practice. *Iranian Journal of Language Teaching Research, 6*(3), 57-77. https://doi.org/10.30466/ijltr.2018.120601

Taras, M. (2008). Summative and formative assessment. *Active Learning in Higher Education*, *9*(2), 172-192. https://doi.org/10.1177/1469787408091655

Topping, K. J. (2009). Peer-assessment. *Theory into Practice*, *48*(1), 20-27. https://doi.org/10.1080/00405840802577569

Turner, C. E., & Purpura, J. E. (2016). Learning-oriented assessment in second and foreign language classrooms. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment*. DeGruyter Mouton.

Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford University Press.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning,* (pp. 53-82). Lawrence Erlbaum Associates.

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing, 21*, 321-331. https://doi.org/10.1016/j.jslw.2012.09.007

Wolters, C. (1999). The relation between high school students' motivational regulation and their use of learning strategies, effort, and classroom performance. *Learning and Individual Differences, 11,* 281-299. https://doi.org/10.1016/S1041-6080(99)80004-1

Zarei, A. A., & Usefi, Z. (2015). The effect of assessment type on EFL learners' goal orientation. *Journal of Language, Linguistics and Literature, 1*(4), 112-119. https://www.aiscience.org/journal/j3l

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance*, (pp. 49-64). Routledge.

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Investigating Cultural Intelligence and Emotional Intelligence as Predictors of Burnout Among Iraqi EFL Teachers

Research Article
pp. 73-93

Reza Bagheri Nevisi*[1]
Zeinab Alasadi[2]

Received: 2020/10/14 Accepted: 2021/02/09

Abstract
The study was an attempt to discover any possible correlation between EFL teachers' burnout and their emotional and cultural intelligence. Furthermore, it aimed at finding out which variable, emotional or cultural intelligence, could be a better predictor of teacher burnout. The study also delved into the possible relationship of gender and Iraqi EFL teachers' burnout and investigated which gender was more affected by burnout. Finally, male and female Iraqi teachers were compared and contrasted with regard to emotional intelligence and cultural intelligence. To this end, three questionnaires- (1) Maslach Burnout Inventory–Educators' Survey (MBI-ES), (2) the Cultural Intelligence Scale (CQS) developed by Van Dyne et al. (2008), and (3) Wang and Law Emotional Intelligence Scale (WLEIS)- were submitted to more than 200 teachers either in person or via email or an already-generated Google Doc link of the three questionnaires. Some teachers remained reluctant to share their responses and refused to fill out the questionnaires, but 164 Iraqi EFL teachers completed all the three questionnaires. The results demonstrated that Iraqi EFL teachers' burnout and their emotional and cultural intelligence correlated significantly but negatively. Both emotional and cultural intelligence could significantly and almost similarly predict teacher burnout. Moreover, teacher burnout and gender were significantly correlated and Iraqi EFL female teachers were more susceptible and vulnerable to teacher burnout. The study implies that emotionally and culturally-aware teachers might be able to take precautionary measures to either mitigate or curb teacher burnout.

keywords: teacher burnout, emotional intelligence, cultural intelligence, EFL teachers, gender

---

*Corresponding Author
[1] Associate Professor, Department of English Language and Literature, University of Qom, Qom, Iran. re.baghery@gmail.com, r.bagherinevisi@qom.ac.ir
[2] MA Graduate, Department of English Language and Literature, University of Qom, Qom, Iran. zalasadi92@yahoo.com

DOI: 10.22051/LGHOR.2021.33516.1383

Introduction

Living happily is largely dependent upon an individual 's gratification with what he or she does. Individuals spend long hours in their workplace and there are many intervening and overlapping variables that could either positively or negatively influence their job performance. For instance, it might be crucial to pay attention to our emotions in managing our behavior and career (Mayer & Salovey, 1997). Furthermore, the ability to manage our feelings and keep them under control could be regarded as an effective means for employers to keep their job stress in check (Mayer, et al., 2004). In the same vein, language teaching is a demanding job among many professions which requires great cultural, emotional, and technical expertise on the part of educators if the profession is to blossom, thrive, and succeed.

Ideally speaking, teachers should be able to delve into target language culture, language teaching methodologies that can be adopted and the way emotional and cultural intelligence impact their overall profession. Emotional intelligence can be described as a set of cognitive abilities, competencies and skills that impact one's capability to succeed in dealing with situational demands (Bar-On, 2001) and cultural intelligence means one's ability to act properly in contexts featured by cultural multiplicity and variety (Van Dyne & et al., 2007). Teaching is not only teaching words, phrases, and grammar but also includes the emotional, cultural, social and contextual variables. Through increasing emotional skills, teacher and students' motivation toward teaching and learning process can be boosted and a helpful educational environment for both can be created and fostered. Accordingly, teacher role will become more and more significant especially if teachers are more culturally and emotionally-ware (Jennings & Greenberg, 2009).

Iraq as a developing country is now moving fact towards pedagogical advancements and research developments in many scientific fields. To boost their current scientific status in the world both with regard to technology and research, Iraqi educational institutions are now offering courses and materials pertaining to Teaching English as a Foreign Language. Such institutions are constantly on the rise and increasingly involved in providing their intended audience with proper learning materials, advice, and guidance as they feel the pressing need to reach for an international readership almost on a daily basis. Although some studies have been done on teacher burnout in EFL settings, this research sets to probe into the relation between cultural intelligence, emotional intelligence and teacher burnout among Iraqi EFL teachers. Given the paucity of such research in that setting, the study seeks to discern whether cultural intelligence and emotional intelligence can be regarded as predictors of teacher burnout among Iraqi EFL teachers. Furthermore, since both Iraqi male and female teachers attended the research, the researchers aimed to find out whether gender could be viewed as a determining factor when it came to the above-mentioned variables.

Literature Review

Burnout

The term "burnout" was first used in the U.S.A. in late 1960s and 1970s by a psychologist named Herbert Freudenberger. Freudenberger (1974) saw that many of his co-workers were emotionally-exhausted. They had no motivation to work after a while. This continued until other psychological and physical signs appeared. Therefore, he created the term “burnout “which was referred to as physical and

psychological exhaustion (Gold & Roth, 1993; Maslach & Schaufeli, 1993; Schaufeli, 2003).

Many definitions have been proposed for the term "burnout". For instance, Maslach (1982) described it as a condition of emotive fatigue, depersonalization, and decreased individual attainment. Pines and Aronson (1988) described burnout as an extreme physical, mental and emotional exhaustion which could be caused by working for long periods of time. Yet another definition was proposed by Wallace and Brinkerhoff (1991) who contended that burnout transpires when workers don't have enough power to respond appropriately to the demands of their job and thus they cannot cope with its negative consequences.

Cultural Intelligence (CQ)

Earley and Ang (2003) describe CQ as one's capability to adjust successfully to the new cultural conditions. To achieve this kind of adaptation, one needs skills and capability different from those used by people within their own cultural context. There are four interrelated factors that lead to the concept of cultural intelligences (Van Dyne et al., 2010). The four aspects of CQ: Cognitive, metacognitive, motivational and behavioral differ from the overall capability to act successfully in divergent cultural contexts (Earley & Ang, 2003). These dimensions reflect a set of adaptive abilities that are of paramount importance for effective personal relations and success in culturally divergent and varying contexts (Van Dyne et al., 2009). Van Dyne et al. (2007) name the following four characteristics for culturally-intelligent individuals. 1. They are good predictors and they can prognosticate what might happen in cross-cultural settings. 2. They are able to apprehend multi-cultural situations. 3. They choose to experience entirely varying and divergent cultural contexts. 4. They can readily adjust themselves to a new environment and modify or alter their verbal or nonverbal behaviors in various contexts.

CQ displays a person's capability to adjust himself or herself to new circumstances and generate proper behavior (Earley & Peterson, 2004). Since not all the new cultural situations will be immediately understood, an important skill to be considered for a culturally-intelligent person is the ability to expect misunderstandings. This is what Triandis (2005) dubs "suspending judgment". In other words, culturally-intelligent individuals possess higher levels of ambiguity tolerance and this allows them to lower their stress levels when placed in a new culture and produce the best possible performance based on a better cultural understanding (Brislin et al., 2006).

Emotional Intelligence

EI was first introduced by Payne (1985) in his doctoral dissertation. The brilliant idea occurred to Salovey one summer when he was painting his home with his friend John Mayer. Salovey investigated emotions and behavior, whereas Mayer examined the association between emotions and thought. Mayer and Salvoery (1997) described emotional intelligence as a group of interconnected skills regarding the capability to apprehend correctly, gauge, and convey feelings and emotive information and the capability to direct, handle, and balance feelings to enhance affective and academic development. Several EI models exist in the literature. Among them, Petrides et al. (2016) in the trait model characterize EI as a set of affect-related personality traits quantifiable with self-reports (Petrides et al., 2007;

Hughes & Evans, 2018). They have presented a short description of the domains where trait EI can be readily applied like organizational, socio-educational, and clinical fields. Two different EI constructs can be distinguished according to the measurement technique that is adopted to operationalize them (self-report techniques as in questionnaires or maximum performance tests as in IQ tests). *Trait EI* deals with affect-related self-conceptions that can be measured through self-report questionnaires, whereas *ability EI* copes with affect-related cognitive capabilities that have to be tested and evaluated through maximum performance tests (Petrides, et al., 2007, 2016). Since the researchers of the present study have turned to a self-report questionnaire, the construct in question belongs to the *Trait* category of emotional intelligence.

EI is comprised of five basic constituents: Self-awareness, self-regulation, motivation, empathy and social skills. Each has its own description. (1) Self-awareness means the ability to understand others' feelings, merits, demerits, wants, and incentives (Goleman, 1995). The second component of EI is self-regulation that makes an individual think about his or her own feelings and enables them not to rush headlong into decisions. The third component is motivation which is one's desire to achieve something. Motivated individuals rarely get exhausted and go to great lengths to produce the best performance and outcome. They are typically optimistic and goal-oriented and successfully accomplish their objectives. The fourth component is empathy that enables an individual to understand others' feelings and form a better relationship with them. The last component is social skills that make individuals to be good persuaders (Goleman, 1995).

Some Relevant Empirical Studies

Many student-related studies on emotional and cultural intelligence have been done in the Iranian EFL setting (e.g. Farooq, 2014; Ghonsooly et al., 2013; Marashi & Zaferanchi, 2010; Mohammad Hosseinpur & Farahani, 2017; Salehzade & Lashkarian, 2015; Shahmohamadi & Hasanzadeh, 2011). However, there are a number of relevant teacher-related studies done on emotional, cultural intelligence, and burnout. Tomic, et al. (2004) delved into the relation between existential fulfillment and teacher burnout. Personal attainment for individuals was gauged through the Existence scale and burnout was gauged via MBI. Findings revealed that there existed a negative association between existential fulfillment and emotional exhaustion and depersonalization but a positive one existed for personal attainment.

Cano-Garcia et al. (2005) examined the significance of personality and contextual factors in teacher burnout. The study aimed at explaining and prognosticating the relationship between levels of teacher burnout, personality factors and some contextual elements. A semi-structured interview, the modified version of the NEO-PI-R and MBI were utilized to gather data. The findings pointed to the significance of personality factors and contextual elements not only in describing levels of teacher burnout, but also in predicting burnout levels in teachers.

Hogan and McKnight (2007) carried out a research to investigate burnout among 76 higher education online instructors. MBI and a researcher-developed demographic survey were applied for data collection. The findings indicated that the instructors enjoyed an average grade on emotive fatigue, a low grade on the personal fulfilment and a relatively high score on the depersonalization subcomponents of the questionnaire.

MacNab and Worthley (2012) probed into the relationship between CQ and individual features. The findings demonstrated that general self-efficacy and CQ were positively correlated. Moreover, the results revealed that the three subcategories of CQ: Behavioral, cognitive, and metacognitive were related to general self-efficacy as well. The study implied that general self-efficacy should be integrated within CQ educational efforts and development as a vital characteristic.

Koçoğlu (2011) probed into the association between emotional intelligence and teacher efficacy. Ninety English language pre-service university teacher attended the research. The results indicated that teachers preferred enjoying efficacy in handling the class over having a fun class. The teachers had the highest score in anxiety tolerance and assistance competencies, but lowest in autonomy and self-consideration concerning emotional intelligence. The results pointed to a positively significant relation between EQ and teacher efficiency.

Bağçeci and Hamamci (2012) examined the relation between teacher burnout and their coping strategies. The Maslach Burnout Inventory, the Problem Solving Inventory, and the Ways of Coping Inventory were utilized. Results demonstrated that a significantly positive relation existed between the helpless coping strategy and teachers' emotional exhaustion. Furthermore, the findings were indicative of the fact that the self-confident coping strategy could be regarded as a predictor of personal accomplishment and was negatively related with it.

Fardinpour and Masoomi (2014) delved into the personality features as likely predicators of teachers' burnout. The results were indicative of the fact that burnout and personality factors were significantly associated with one another. The study concluded that personality features could predict levels of teacher burnout and probably enable educators to make necessary precautionary measures to either mitigate fatigue, depression, and burnout or keep such negative feelings in check to a great extent.

Mérida-López and Extremera (2017) studied the relation between EI and teacher burnout. The research aimed at implementing a systematic EI and teacher burnout. MEDLINE, PSYCinfo and Scopus databases were thoroughly investigated and 13 qualified scholarly papers were picked. The results showed that the two were negatively associated.

Esmaili et al. (2018) investigated EFL teachers' emotional intelligence and burnout in a language institute. Sixty-three language teachers from 11 private language institutes attended the research. Bar-On 's EQ questionnaire and Maslach Burnout Inventory were employed. The findings demonstrated that teacher's emotional intelligence and the three subcategories of burnout (emotional exhaustion, depersonalization and accomplishment) were significantly related. Although some studies have been carried out in EFL contexts with regard to teacher burnout, personality factors, teacher empowerment and teacher reflectivity, the relationships between cultural intelligence and Iraqi EFL teachers' burnout on the one hand, and emotional intelligence and their likely burnout on the other hand, have remained largely unexplored. To attain the above-stated purposes of the research, the researchers formulated these questions:

1. Is there any significant correlation between Iraqi EFL teachers' emotional intelligence and burnout?

2. Is there any significant association between Iraqi EFL teachers' cultural intelligence and burnout?

3. Which variable (emotional intelligence or cultural intelligence) can be regarded as a better predictor of teacher burnout among Iraqi EFL teachers?

4. Are there any significant differences between male and female Iraqi EFL teachers with regard to burnout, emotional intelligence, and cultural intelligence?

Methodology

Participants

A hundred and sixty-four Iraqi teachers including 97 females and 67 males participated in the research. Their age ranged from 23 to 55 years old. Their teaching experience ranged from 1 to 25 years. Moreover, Arabic language accounted for all the participants' first language. The respondents were either met in person or electronically contacted from various schools, universities and language institutes from different cities of Iraq (e.g. Basra, Baghdad, Karbala, Najaf, and Dhiqar). Ninety-seven teachers accounting for 59.14% of all the participants were TEFL graduates, 46 teachers accounting for 28.04% of all the participants were English Literature graduates, and 21 teachers accounting for 12.80% of all the participants were English Translation graduates. Concerning the geographical location of the teachers, 25 of the participants took part in the study from Basra, 63 from Baghdad, 26 from Karbala, 27 from Najaf, 17 from Dhiqar, and 6 from Kufa.

Table 1

Demographic Information of the Participants

Average Age	Average Teaching Experience	Academic Degree	Institution
23-34 (N = 19)	1-7 Years (N = 17)	B.A. (N = 80)	Schools (N = 89)
34-44 (N = 68)	7-12 Years (N = 87)	M.A. (N = 66)	Institutes (N = 58)
44-55 (N = 77)	12-25 Years (N = 60)	Ph.D. (N = 18)	Universities (N = 17)

Instruments

As for the instruments, the researchers employed the original versions in English and made minor modifications wherever an indication was deemed necessary to further specify that the questionnaires were being particularly used for the teaching profession. Furthermore, the participants were provided with no further assistance such as explanations or translations from the researchers. The three instruments utilized in the study are as follows:

Maslach Burnout Inventory –Educators' Survey (MBI-ES). The questionnaire consists of 22 self-report items and includes three subcomponents of Emotional Exhaustion (EE), Depersonalization (DP) and Personal Accomplishment (PA). The first two subcomponents: (EE & DP) contain seven items each and PA includes eight items pertaining to feelings of personal attainment and success. All the participants were kindly requested to go through the relevant sentences for each subcomponent and provide their responses ranging from never accounting for 0 to everyday accounting for 6 on a 7-point Likert scale. A high score in the first two sections and a low score in the last section might be an indication of burnout.

The Cultural Intelligence Scale (CQS) Developed by Van Dyne et al. (2008). It is composed of 20 items which includes four metacognitive, six cognitive, five motivational and five behavioral items. This instrument enjoys high reliability i.e. a = 0.86 and the reliability of its cognitive, meta-cognitive, motivational, and behavioral subcomponents stand at 0.81, 0.82, 0.64, 0.72 respectively. The metacognitive component is concerned with a person's capability to manage diverse

cultural systems and entails planning, goal-setting, monitoring performances, being cognizant of cross-cultural variations in different settings, and evaluating the outcomes of the situations. The cognitive component deals with a person's knowledge about the host culture and it consists of business, interpersonal, and socio-linguistics knowledge of that specific culture. The motivational component is associated with an individual's capability to draw attention towards knowing how to properly function in various cross-cultural settings and integrates intrinsic and extrinsic motivation with self-efficacy. The behavioral dimension of cultural intelligence is linked with one's capability to appropriately adopt verbal and nonverbal behaviors of the host culture and is representative and reflective of the suitable manners at the host culture (Van Dyne, et al., 2012).

Wang and Law Emotional Intelligence Scale (WLEIS). The Bar-On 's emotional intelligence inventory (EQ-i) comprising 5 factors and 15 elements is a commonly-used EQ questionnaire in the literature. The initial account had 133 questions. Later Bar-On (1997) amended the questionnaire and decreased its size to 117 items. Due to the feasibility issue, the challenges inherent in administering questionnaires with too many items, and the fact that the researchers desired and asked for the participants' active collaboration, cooperation, and full engagement with the research project, the researchers opted for this specific and already-validated questionnaire to either mitigate or curb the cognitive load imposed upon the respondents' working-memory and glean their full cooperation as well. This questionnaire is a 16-item self-report instrument. The respondents had to respond to the items on a 7-point Likert Scale ranging from strongly disagree to strongly agree.

KR-21 Reliability Indices for the Instruments

Table 2 illustrates the descriptive statistics and KR-21 reliability indices for the burnout, and emotional and cultural intelligence questionnaires. The results demonstrated that the reliability indices for the burnout, emotional and cultural intelligence questionnaires were .80, .71 and .72 respectively.

Table 2
KR-21 Reliability Indices for Burnout, Emotional and Cultural Intelligence

	N	Min	Max	Mean	Std. Deviation	Variance	KR-21
Burnout	164	59	121	87.15	13.691	187.451	.80
Emotional Intelligence	164	36	83	60.00	9.659	93.301	.71
Cultural Intelligence	164	42	96	68.95	11.085	122.887	.72

Data Collection Procedure

The three questionnaires were submitted to more than two-hundred Iraqi EFL teachers either in person or electronically through their emails or an already-designed Google Doc link of the questionnaires for those who were physically distant from the researchers. Some teachers remained reluctant to share their responses and refused to fill out the questionnaires and 164 Iraqi EFL teachers completed all the three questionnaires. All the respondents were ensured that the information they had provided would remain confidential and wouldn't be released without their due consent.

Data Analysis

First, there was a need to check if the present data suffered from any univariate outliers or not. To this end, the researchers measured the Z scores. Then

the Mahalanobis Distances (MD) were computed in order to probe any multivariate outliers. The next step was to check the assumption of normality through skewness and kurtosis indices and their ratios over the standard errors. Since the absolute value of ratios of skewness and kurtosis were lower than 1.96 for all variables, the research questions were analyzed through parametric Pearson correlation and linear regression. In order to respond to the first and second question, the researchers computed Pearson correlation. This was to probe any significant relation between Iraqi EFL teachers' emotional and cultural intelligence and their burnout. A linear regression using backward method was run to probe which of the two intelligences; i.e. emotional or cultural, was a better predictor of burnout. Since the third research question required which one of the intelligences was a better predictor of burnout, the two regression coefficients were compared for any significant differences using the online calculator. Moreover, the assumptions of linearity of relationship between the two variables in each research questions and homoscedasticity (homogeneity of variances) were probed through Scatter Plot. At the end, an independent-samples t-test was conducted to compare the male and female groups' means on burnout, emotional intelligence and cultural intelligence separately.

Results

Table 3 illustrates the descriptive statistics for the standardized scores (Z-scores) of burnout, cultural and emotional intelligences. Since none of the Z-scores were higher than +/- 3, it was concluded that the data did not suffer from any univariate outliers.

Table 3
Descriptive Statistics for Standardized Scores of Burnout, Emotional and Cultural Intelligences

Z-Score	N	Minimum	Maximum	Mean	Std. Deviation
Burnout	164	-2.055	2.472	.000	1.00
Emotional Intelligence	164	-2.484	2.381	.000	1.00
Cultural Intelligence	164	-2.430	2.440	.000	1.00

The Mahalanobis Distances (MD) were computed in order to probe any multivariate outliers. The computed MD's, as displayed in Table 4, were compared against the critical value of chi-square at .001 levels for three variables, i.e. 16.26. Table 4 illustrates the descriptive statistics for the MD. Since maximum MD of 9.85 was lower than critical value of chi-square at 3 for the three variables; i.e. 16.26. The data did not suffer from any multivariate outliers.

Table 4
Descriptive Statistics of Mahalanobis Distances; Testing Multivariate Outliers

	N	Minimum	Maximum	Mean	Std. Deviation
Mahalanobis Distance	164	.101	9.855	2.981	2.187

The assumption of normality was checked through skewness and kurtosis indices and their ratios over the standard errors (Table 5). The absolute value of ratios of skewness and kurtosis were lower than 1.96 for all variables. That was why the research questions were analyzed using parametric Pearson correlation and linear regression.

Table 5
Descriptive Statistics; Testing Normality of Data

	N	Skewness			Kurtosis		
	Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
Burnout	164	.348	.190	1.83	-.469	.377	-1.24
Emotional Intelligence	164	-.010	.190	-.052	-.149	.377	-.395
Cultural Intelligence	164	.101	.190	.531	-.393	.377	-1.04

Exploring Question 1

The first question intended to discern whether there was any significant relation between Iraqi instructors' EI and burnout. Pearson correlation was computed in order to probe into any possible association. According to the findings illustrated in Table 6, [r (162) = - .567, denoting a big effect size, p = .000], it could be said that Iraqi instructors' EI and their burnout significantly but negatively correlated with one another. The negative and significant correlation between burnout and emotional intelligence indicated that if Iraqi EFL teachers' burnout increased their emotional intelligence decreased.

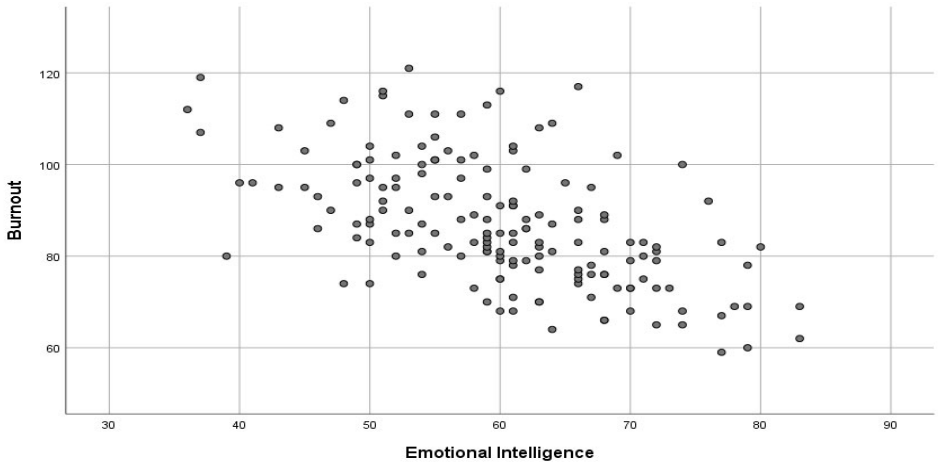
Table 6
Pearson Correlation between Emotional Intelligence and Burnout

		Emotional Intelligence
Burnout	Pearson Correlation	-.567**
	Sig. (2-tailed)	.000
	N	164

**. Correlation is significant at the 0.01 level (2-tailed).

As mentioned earlier, Pearson correlation, besides the assumptions of lack of univariate and multivariate outliers and normality, assumes linearity of relation between the two variables and homoscedasticity (homogeneity of variances). These two assumption were probed through Scatter Plot 1. Since the spread of dots did not show any rising-and-falling pattern, it was concluded that the assumption of linearity was retained. The spread of dots also did not form any funnel shape, i.e. wide at one end and narrow at the other end, it was concluded that the assumption of homoscedasticity was retained.

Scatter Plot 1
Relationship between Burnout and Emotional Intelligence



Exploring Question 2

The second question examined whether there was any statistically meaningful relation between Iraqi EFL teachers' cultural intelligence and burnout. Pearson correlation was computed in order to probe that relation. According to the findings illustrated in Table 7, [r (162) = - .553, depicting an enormous effect size, p = .000], it could be said that a significantly negative and large association existed between Iraqi instructors' CQ and their burnout. The negative and significant correlation between burnout and cultural intelligence indicated that if Iraqi EFL teachers' burnout increased their emotional intelligence decreased.

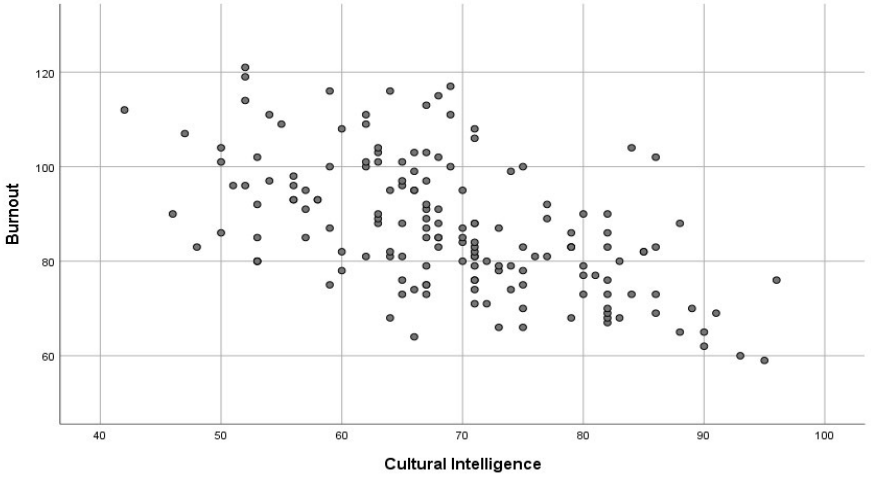
Table 7
Pearson Correlation between Cultural Intelligence and Burnout

		Cultural Intelligence
Burnout	Pearson Correlation	-.553**
	Sig. (2-tailed)	.000
	N	164

**. Correlation is significant at the 0.01 level (2-tailed).

The assumptions of linearity of relationship between the two variables and homoscedasticity (homogeneity of variances) were probed through Scatter Plot 2. Since the spread of dots did not show any rising-and-falling pattern, it was concluded that the assumption of linearity was retained. The spread of dots also did not form any funnel shape, i.e. wide at one end and narrow at the other end, it was concluded that the assumption of homoscedasticity was retained.

Scatter Plot 2
Relationship between Burnout and Cultural Intelligence



Exploring Question 3

The third question intended to determine which variable (emotional intelligence or cultural intelligence) could be regarded as a better predictor of teacher burnout among Iraqi EFL teachers. A linear regression using backward method was run to probe which of the two intelligences; i.e. emotional or cultural,

was a better predictor of burnout. The Backward method was selected because as noted by Field (2018), the Forward and Stepwise methods should be avoided. Table 8 displays the findings of the regression analysis. The findings indicated that both emotional and cultural intelligences predicted 36.2 percent of Iraqi EFL teachers' burnout (R = .602, R^2 = .362).

Table 8
Model Summary[b]

Model	R	R Square	Adjusted R Square
1	.602[a]	.362	.354

a. Predictors: (Constant), Cultural Intelligence, Emotional Intelligence
b. Dependent Variable: Burnout

Table 9 displays the findings of the ANOVA test of significance of regression model. The results (F (2, 161) = 45.65, p = .000, partial η^2 = .221 depicting a big effect size) demonstrated that the regression model enjoyed statistical significance. In other words; emotional and cultural intelligences significantly predicted burnout.

Table 9
ANOVAa Test of Significance of Regression Model

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	11057.266	2	5528.633	45.653	.000b
	Residual	19497.222	161	121.101		
	Total	30554.488	163			

a. Dependent Variable: Burnout
b. Predictors: (Constant), Cultural Intelligence, Emotional Intelligence

Table 10 exhibits the findings of regression coefficients. The two sets of regression coefficients were provided; standardized (beta) and unstandardized (b) values. The standardized regression coefficients show the amount of change in dependent variable (burnout) due to one standard deviation change in the predictor. For example, the beta value for emotional intelligence was -.350. That is to say, if emotional intelligence increases one standard deviation, burnout decreases .350 standard deviations. The unstandardized regression coefficients (b) are interpreted in terms of the unit of measurement used to measure the variables. For example, the b-value for cultural intelligence was -.365. That is to say, if cultural intelligence increases one unit, burnout decreases .365 units.

Table 10
Regression Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
Model		B	Std. Error	Beta		
1	(Constant)	142.051	5.813		24.436	.000
	Emotional Intelligence	-.496	.132	-.350	-3.759	.000
	Cultural Intelligence	-.365	.115	-.295	-3.173	.002

a. Dependent Variable: Burnout

According to Table 10, it can be said that both emotional intelligence (b = -.496, Beta = -.350, t = -3.75, p = .000) and cultural intelligence (b = -.365, Beta = -.295, t = -3.17, p = .002) significantly predicted burnout. However, the third research question required which one of the intelligences was a better predictor of burnout.

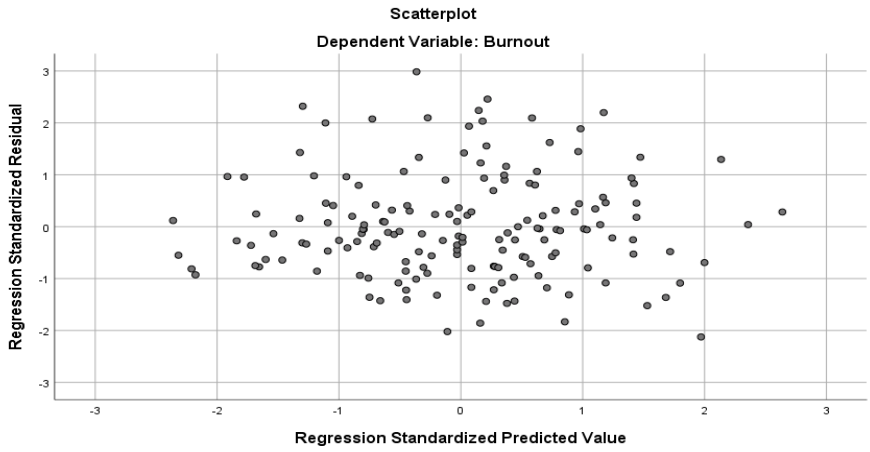
The two regression coefficients were compared for any significant differences using the online calculator. The results (Table 11) indicated that there was not any (t = .748 (324), p = .454) significant difference between the two predictors. That is to say; both emotional and cultural intelligences significantly predicted burnout.

Table 11
Comparing two Regression Coefficients

Intelligences	b-values (Slopes)	Standard Errors	t-value	Df	p-value
Emotional	-.496	.132	.748	324	.454
Cultural	-.365	.115			

The assumptions of linearity of relationship between the two variables and homoscedasticity (homogeneity of variances) were probed through Scatter Plot 3. Since the spread of dots did not show any rising-and-falling pattern, it was concluded that the assumption of linearity was retained. The spread of dots also did not form any funnel shape, i.e. wide at one end and narrow at the other end, it was concluded that the assumption of homoscedasticity was retained.

Scatter Plot 3
Testing Linearity and Homoscedasticity of Regression model



Exploring Question 4

An independent-samples t-test was conducted to examine the male and female instructors' means on burnout. Table 12 displays the findings of the descriptive statistics for the two groups on burnout. The results indicated that the female Iraqi teachers (M = 94.26, SD = 11.48) had a higher burnout than the male teachers (M = 76.85, SD = 9.43).

Table 12
Descriptive Statistics of Burnout by Gender

	Group	N	Mean	Std. Deviation	Std. Error Mean
Burnout	Female	97	94.26	11.487	1.166
	Male	67	76.85	9.439	1.153

Table 13 displays the findings of the independent-samples t-test. The results (t (157) = 10.61, p = .000, 95 % CI [14.16, 20.64], r = .646 signifying a considerable effect size) demonstrated that the female Iraqi EFL teachers had a significant higher burnout than the male teachers. Homogeneity of variances assumption was not maintained. As illustrated in Table 13, the significant findings of the Levene's test (F = 5.65, p = .019) revealed that the two groups' variances on burnout differed significantly. That was why the second row of Table 13; i.e. "Equal variances not assumed" was reported.

Table 13
Independent-Samples t-test; Burnout by Gender

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	5.657	.019	10.241	162	.000	17.407	1.700	14.051	20.763
Equal variances not assumed			10.613	157.082	.000	17.407	1.640	14.167	20.647

An independent-samples t-test was conducted to juxtapose the male and female instructors' means on emotional intelligence. Table 14 shows the descriptive statistics for the two groups on emotional intelligence. The results indicated that the male Iraqi teachers (M = 65.30, SD = 8.46) had a higher emotional intelligence than their female counterparts (M = 56.34, SD = 8.71).

Table 14
Descriptive Statistics of Emotional Intelligence by Gender

	Group	N	Mean	Std. Deviation	Std. Error Mean
Emotional Intelligence	Male	67	65.30	8.466	1.034
	Female	97	56.34	8.719	.885

Table 15 exhibits the findings of the independent-samples t-test. The results (t (162) = 6.54, p = .000, 95 % CI [6.25, 11.66], r = .457 depicting a modest to big effect size) demonstrated that the male Iraqi EFL instructors enjoyed a significantly higher emotional intelligence compared to their female counterparts.

Table 15
Independent-Samples t-test; Emotional Intelligence by Gender

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.011	.916	6.545	162	.000	8.958	1.369	6.255	11.661
Equal variances not assumed			6.580	144.728	.000	8.958	1.361	6.268	11.649

The homogeneity of variances assumption was kept. As shown in Table 15, the non-significant findings of the Levene's test (F = .011, p = .916) demonstrated that the two groups' variances on emotional intelligence did not differ significantly from one another. That was why the second row of Table 15; i.e. "Equal variances assumed" was reported.

Table 16 displays the descriptive statistics for the two group on cultural intelligence. The results indicated that the male Iraqi teachers (M = 74.79, SD = 9.18) had a higher cultural intelligence than their female counterparts (M = 64.91, SD = 10.50).

Table 16
Descriptive Statistics of Cultural Intelligence by Gender

	Group	N	Mean	Std. Deviation	Std. Error Mean
Cultural Intelligence	Male	67	74.79	9.189	1.123
	Female	97	64.91	10.501	1.066

Table 17 displays findings of the independent-samples t-test. The results (t (162) = 6.23, p = .000, 95 % CI [6.75, 13.01], r = .440 signifying a modest to large effect size) indicated that the male Iraqi EFL teachers had a significant higher cultural intelligence than the female teachers. The homogeneity of variances assumption was maintained. As illustrated in Table 17, the non-significant findings of the Levene's test (F = 5.65, p = .019) demonstrated that the two groups' variances on cultural intelligence did not differ significantly from one another. That was why the first row of Table 17; i.e. "Equal variances assumed" was reported.

Table 17
Independent-Samples t-test; Cultural Intelligence by Gender

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.489	.485	6.230	162	.000	9.884	1.587	6.751	13.017
Equal variances not assumed			6.384	153.117	.000	9.884	1.548	6.825	12.943

Discussion

The present study aimed at finding out the relation between Iraqi instructors' burnout and their emotional and cultural intelligence. The study also attempted to discern whether emotional and cultural intelligence could predict teacher burnout among Iraqi EFL teachers. Furthermore, male and female Iraqi teachers were compared and contrasted with regard to teacher burnout, emotional intelligence, and cultural intelligence. It was concluded that teacher burnout significantly related with both emotional and cultural intelligence. In other words, the higher a teacher emotional intelligence and cultural intelligence were, the lower his or her burnout could be. Both emotional and cultural intelligence could predict levels of teacher burnout. Iraqi female teachers were more impacted by teacher burnout and consequently were more emotionally and culturally vulnerable than their male colleagues.

The first research question delved into the possible relation between Iraqi instructors' burnout and their emotional intelligence. The findings indicated a negative and significant association between the two. As emotional intelligence increased, levels of teacher burnout decreased in the participants. This finding could reasonably be justified on the grounds that those who are more capable of managing their emotions under tough and difficult circumstances might suffer less from teacher burnout and be less prone to job dissatisfaction, fatigue or even depression. This concurs with Esmaili et al. (2018) and Mérida-López and Extremera (2017) who found a significant relation between instructors' EI and the three dimensions of burnout.

Being emotionally-cognizant is an important component when it comes to making proper, informed, and appropriate decisions in chaotic or critical conditions. Typically, when an individual is under pressure and emotions are mixed and abound, the decision-making process will be negatively influenced and the one who suffers as a result is the decision-maker. Teaching can be regarded as both a highly demanding and challenging profession simultaneously. There are times when teachers feel exhausted or challenged by the daunting tasks they encounter inside and outside the classroom environment, they feel pressured by the heavy responsibilities on their shoulder and they will have to keep their emotions in check to make the right decision. When someone is not emotionally-balanced, the work he or she is involved in doing will surely be adversely affected. In order for a teacher to succeed in his or her profession or career, he or she will have to be mentally prepared and emotionally ready to face the challenges. Otherwise, teacher burnout will be a possible consequence to suffer from. Therefore, it behooves anyone involved in the teaching profession to pay due and sufficient attention to the emotional dimension and raise such an awareness in them accordingly. This is in line with Bağçeci and Hamamci (2012) and Koçoğlu (2011) who found EQ and teacher's efficacy were positively and significantly correlated.

The second research question probed into the possible correlation between Iraqi teachers' burnout and their cultural intelligence. The results were indicative of a significant relationship between the two. In other words, the more culturally-aware a person was, the less likely he or she could be to suffer from high levels of teacher burnout. Such a result could be reasonably justified on the grounds that many teachers embark upon the process of teaching without a sound and proper knowledge, awareness, and understanding of the cultural issues. To be less prone to

high levels of teacher burnout, teachers will have to be appropriately cognizant of the second or foreign language culture. Such an understanding could enable the teacher to predict the challenges and find proper solutions to resolve them as they emerge. Cultural issues play a pivotal role when it comes to a teacher's job, sense of accomplishment and success. If a teacher is not culturally-aware, he or she might be more vulnerable to teacher burnout because such confusion and bewilderment on the part of language teacher might lead to early disappointment and frustration and make the teacher give up early on and prematurely. A contrastive analysis and analogy of the first language and second language culture can be deemed necessary if teachers are to succeed and thrive in their profession. Cultural unfamiliarity with L2 can be considered an impediment to the success of a teacher. This is consistent with Kelidbari et al. (2012) who did a research on the association between cultural intelligence and job performance. The results showed that CQs were good predictors for job performance.

The last research question addressed the relation between Iraqi EFL teachers' burnout and gender and examined which gender was more susceptible to and affected by teacher burnout. The findings revealed that teacher burnout and gender significantly correlated. In the same vein, it was found that Iraqi EFL female teachers were more likely to suffer from teacher burnout and were more affected by burnout than their male counterparts. We all know the fact that men and women are created differently and each possess their God-given peculiarities, qualities and specialties. It is commonly stated that women are more emotional and it is almost taken for granted that they should not be placed and appointed for sensitive positions such as judging. It is true that women are full of emotions, but it remains to be seen whether they can rightly and properly handle and manage their emotions. The results of this research project demonstrated that Iraqi EFL female teachers might be regarded as emotionally more vulnerable and susceptible than their male counterparts. This stands to reason that female teachers suffered more from job dissatisfaction, failure, burnout, and frustration. Another possible explanation for the higher female vulnerability and susceptibility to teacher burnout could be ascribed to the setting or context in which they have grown up. Iraq is located in a volatile region in the middle-east where women are almost regarded as second-class citizens.

In a male-dominated country like Iraq where women are marginalized, not given priority number one or given secondary roles, females might feel suppressed, pressured, marginalized, and underprivileged. Such feelings of inadequacy, inferiority, and marginalization could have contributed to and undermined their overall feelings of job satisfaction and success. If Iraqi women are to thrive and succeed in their teaching profession and be less prone to teacher burnout, they will have to be given more opportunities to unleash their full potentials and their feelings, desires, and attitudes should be properly and adequately taken care of. If Iraqi female teachers feel that they are being cared as much as their male colleagues and the work they do is appreciated as much as Iraqi male teachers, they might be less likely to suffer from high levels of teacher burnout. In fact, such depreciation, apathy, and disregard for their feelings, desires, and tendencies could have been a contributing factor when it comes to teacher burnout

Another important point be born in mind is the amount of salary Iraqi female teachers earn in comparison with their male mates. In other words, a likely justification for the higher levels of teacher burnout among Iraqi female teachers could be ascribed to the fact that they are given lower salaries for the same job

compared to the male teachers. The instrumental and extrinsic motivation inherent in the teaching profession can be regarded as a crucial and determining factor. Such undesired discrimination between Iraqi male and female teachers concerning the amount of money they earn could be considered a likely justification for higher levels of teacher burnout among female Iraqi teachers. The economic aspect of teaching or the materialistic outlook could have impacted female Iraqi EFL teachers' levels of burnout.

Cultural diversities abound between English and Arabic. Such diversities, disparities, discrepancies, and differences between the two languages could have accounted for a low and insufficient cultural awareness and understanding on the Iraqi EFL teachers' part. In fact, when it comes to cultural understanding and awareness, bias and prejudice come into play their part. In a highly-religious Islamic and traditional milieu in Iraq where women have to observe specific cultural norms and regulations, such sharp contrasts between English and Arabic language can be more noticeable and conspicuous. Female Iraqi teachers could certainly be more constrained by their native language sociocultural and religious norms, traditions, and rules. Any violation of such norms and regulations could result in social isolation and exclusion. Iraqi EFL female teachers were less culturally-aware because they might have been highly inundated and engrossed in their own cultural values and traditions. The fact that the Iraqi female teachers could be judged, punished or even isolated for any deviations from their native language cultural norms might be another potential explanation for their higher burnout and lower cultural awareness as well. Feelings of success and gratification among teachers could entail a true and suitable understanding of the cultural likenesses and dissimilarities between the first language and second language. Any digression and deviation from the cultural norms of the first language by females could put them at great risk since bias and prejudice against women in a highly traditional and religious country like Iraq are commonplace. Such fear of exclusion, isolation, and bias against women could have prevented them from gaining a proper cultural awareness.

Conclusion

Generally speaking, teachers will be more vulnerable and susceptible to higher levels of burnout, mental exhaustion, physical fatigue, disappointment, job failure and dissatisfaction if they are less emotionally and culturally-aware. EI does not necessarily mean being emotional but rather it entails a true understanding of one's own feelings, desires, attitudes, and tendencies and being able to manage one's emotions once placed under chaotic and critical circumstances. It includes making the right decision at the right time. One might be a very emotional person but unable to cope with his or her feelings and act properly. Moreover, a suitable cultural awareness and consciousness of both source and target language will assist the teacher to cope with difficulties and challenges, make contingency plans, and find practical and viable solutions for the anticipated problems.

In the teaching world, those teachers who are emotionally intelligent, try their best to develop and enhance their relationship with their students and co-workers (Kremenitzer & Miller, 2008). For language learners, those who have higher EI can more successfully control their stress, and act positively during challenges (Pishghadam, 2009). More specifically, both negative and positive emotions influence the learning quality. Negative emotions can reduce language

learning quality and positive emotion can facilitate language learning (Shau et al., 2013).

A teacher could increase his or her chances of success and job satisfaction, provided that he or she knows how to manage their emotions and how to take into account the cultural issues. Practical workshops can be held to raise teachers' awareness and consciousness about the ways that both emotional and cultural issues can be properly dealt with. The consciousness-raising aspect of such workshops will assist and enable teachers to strike a balance between the desired levels of cultural and emotional intelligence in order not to suffer from the undesirable outcomes of teacher burnout.

The study has a number of implications for the people involved within the field of language teaching. Teachers might benefit from the findings of this research project as they are made aware and cognizant of two important predictors of teacher burnout: Emotional and cultural intelligence. First and foremost, if cultural intelligence can be regarded as one possible predictor of teacher burnout, teachers and practitioners in the field can take precautionary and necessary measures to not only raise teachers' awareness about what it entails, but also the way it can be rectified and boosted. Second, emotionally-aware teachers can also take predetermined steps to prevent teacher burnout or keep it in check to the extent possible. Finally, apart from the consciousness-raising aspect of this research project, it can also serve a preventative function to mitigate teacher burnout.

The present study has its own limitations. First, it is a descriptive and questionnaire-based one. Triangulation of data could be done with the inclusion of other instruments and fewer participants such as interview and observation to increase the overall validity of the research. Due to the feasibility challenges inherent in the administration of questionnaires with too many items, the researcher employed and capitalized on the most well-known and commonly-used ones with fewer items. The study also delimited itself to Iraqi EFL teachers which could have undermined the generalizability of the findings.

Future studies can be done with fewer subjects with instruments other than the utilized questionnaires in this study. Triangulation of data with the inclusion of instruments like interview and observation can be recommended to boost the overall validity of the research. To boost the generalizability of the findings, future studies can also be conducted in ESL contexts to see how similar or different the results might be from the EFL setting so that these two contexts can be compared and contrasted. Variables other than emotional and cultural intelligence that might have a bearing on the teacher burnout could be investigated as well.

References

Bağçeci, B., & Hamamci, Z. (2012). An investigation into the relationship between burnout and coping strategies among teachers in Turkey. *International Journal of Humanities and Social Science*, *2*(12), 67-72. https://doi.org/10.1.1.1063.7243

Bar-On, R. (2001). Emotional intelligence and self-actualization. In J. Ciarrochi, J. Forgas, & J. D. Mayer (Eds.), *Emotional intelligence in everyday life: A scientific inquiry* (pp. 82-97). Psychology Press. https://psycnet.apa.org/record/2001-05487-005

Brislin, R., Worthley, R., & Macnab, B. (2006). Cultural intelligence: Understanding behaviors that serve people's goals. *Group & Organization Management*, *31*(1), 40-55. https://doi.org/10.1177/1059601105275262

Cano-Garcia, F. J., Padilla-Munoz, E. M., & Carrasco-Ortiz, M. A. (2005). Personality and contextual variables in teacher burnout. *Personality and Individual Differences, 38*(4), 929-940. https://doi.org/10.1016/j.paid.2004.06.018

Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures*. Stanford University Press. https://www.sup.org/books/title/?id=3184

Earley, P. C., & Peterson, R. S. (2004). The elusive cultural chameleon: Cultural intelligence as a new approach to intercultural training for the global manager. *Academy of Management Learning & Education*, *3*(1), 100-115. https://www.jstor.org/stable/40214236?seq=1

Esmaili, R., Khojaste, L., & Kafipour, R. (2018). The relationship between emotional intelligence and burnout among EFL teachers teaching at private institutions. *Social Sciences and Humanities*, *26*(3), 1595-1616. https://doi.org/987-363-853-176.

Fardinpour, F., & Masoomi, M. (2014). Personality traits as predictor of EFL teachers' burnout. *International Journal of Language Learning and Applied Linguistics World (IJLLALW),7*(3), 504-512.

Farooq, M., U. (2014). Emotional intelligence and language competence: A case study of the English language learners at Taif University English language center. *Studies in Literature and Language*, *8*(1), 6-9. https://doi.org/10.3968/j.sll.1923156320140801.4262

Field, A. (2018). *Discovering statistics using IBM SPSS*: *Statistics for statistics.* SAGE Publications. https://us.sagepub.com/en-us/nam/discovering-statistics-using-ibm-spss-statistics /book260423

Freudenberger, H., J. (1974). Staff burnout. *Journal of Social Issues*, *30*(1), 159-165. https://doi.org/10.1111/j.1540-4560.1974.tb00706.x

Ghonsooly, B., Sharififar, M., Sistani, S. R., & Ghahari, S. (2013). Cultural intelligence in foreign language learning contexts. *Cultus*, *47*(6), 47-68. http://www.cultusjournal.com/files/Archives/ghonsooly_et_al.pdf

Gold, Y., & Roth, R. A. (1993). *Teachers managing stress and preventing burnout: The professional health solution*. Routledge. https://lib.ugent.be/en/catalog/rug01:000406647

Goleman, D. (1995). *Emotional intelligence*. Bantam books. https://psycnet.apa.org/record/1995-98387-000

Hogan, R. L., & McKnight, M. A. (2007). Exploring burnout among university instructors: An initial investigation. *Internet and Higher Education*, *10*(2),117-124. https://doi.org/10.1016/j.iheduc.2007.03.001

Kelidbari, H. R. R., Dizgah, M. R., & Jourshari, P. R. (2012), The relationship between cultural intelligence and job performance of operational staff in Ports. *Journal of*

Basic and Applied Scientific Research, *2*(6), 6133-6138. https://www.textroad.com/pdf/JBASR/J.%20Basic.%20Appl.%20Sci.%20Res.,%202(6)6133-6138,%202012.pdf

Koçoğlu, Z. (2011). Emotional intelligence and teacher efficacy: A study of Turkish EFL pre-service teachers. *Teacher Development*, *15*(4), 471-484. https://doi.org/10.1080/13664530.2011.642647

Kremenitzer, J. P., & Miller, R. (2008). Are you a highly qualified, emotionally intelligent early childhood educator? *Young Children, 63*(4), 106-112. https://eric.ed.gov/?id=EJ819266

Jennings, P.A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, *79*(1), 491-525. https://doi.org/10.3102/0034654308325693

MacNab, B., & Worthley, R. (2012). Individual characteristics as predictors of cultural intelligence development: The relevance of self-efficacy. *International Journal of Intercultural Relations*, *36*(1), 62-71. 10.1016/j.ijintrel.2010.12.001

Marashi, H., & Zaferanchi, Z. (2010). The relationship between EFL teachers' emotional intelligence and their effectiveness in managing young learners' classrooms. *Journal of English Language Studies*, *1*(4), 85-11.

Maslach, C. (1982). *Burnout: The cost of caring.* Prentice-Hall. https://doi.org/10.1176/ps.34.7.650

Maslach, C., & Schaufeli, W. B. (1993). *Historical and conceptual development of burnout.* In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Series in applied psychology: Social issues and questions. Professional burnout: Recent developments in theory and research* (pp. 1-16). Taylor & Francis. https://psycnet.apa.org/record/1993-97794-001

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence*: *Educational implications* (pp. 3-31). Basic Books. https://scholars.unh.edu/psych_facpub/422/

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *15*(3), 197-215. https://doi.org/10.1207/s15327965pli1503_02

Mérida-López, S., & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research*, *85*(5), 121-130. https://doi.org/10.10.16/j.ijer.2017.07.006

Mohammad Hosseinpur, R., & Sarbandi Farahani, M. (2017). Cultural intelligence, cultural identity, and EFL learners' use of politeness strategies. *Journal of Modern Research in English Language Studies, 4*(4), 27-45. https://doi.org/10.30479/elt.2017.1453

Payne, W. L. (1985). *A study of emotion*: *Developing emotional intelligence*; *self-integration; relating to fear, pain and desire* [Unpublished doctoral dissertation]. The Union for Experimenting Colleges and Universities, Ohio. https://philpapers.org/rec/PAYASO

Petrides, K. V., Mikolajczak, M., Mavroveli, S., Sanchez-Ruiz, M. J., Furnham, A., & Pérez-González, J. C. (2016). Developments in trait emotional intelligence research. *Emotion Review, 8*(4),335-341. https://doi.org/10.1177/1754073916650493

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology, 98*(2), 273-289. https://pubmed.ncbi.nlm.nih.gov/17456273/

Pines, A., & Aronson, E. (1988). *Career burnout: Causes and cures*. The Free Press. https://psycnet.apa.org/record/1988-98289-000

Pishghadam, R. (2009). A quantitative analysis of the relationship between emotional

intelligence and foreign language learning. *Electronic Journal of Foreign Language Teaching*, *6*(1), 31-41.

Salehzade, M., & Lashkarian, A. (2015). The relationship between emotional intelligence and verbal intelligence in Iranian EFL learners. *International Journal of Educational Investigations*, *2*(6), 88-95. http://www.ijeionline.cm/attachments/article/43/IJEI_Vol.2_No.6_2015-6-09.pdf

Schaufeli, W. B. (2003). Past performance and future perspectives of burnout research. *South African Journal of Industrial and Organizational Psychology*, *29*(4), 1-15. https://doi.org/10.4102/sajip.v29i4.127

Shahmohamadi, F., & Hasanzadeh, R. (2011). Emotional intelligence and its predictive power in Iranian foreign language learners' language achievement [Paper presentation]. International Conference on Social Science and Humanity, ACSIT Press, Singapore. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.477.5076&rep=rep1&type=pdf

Shau, K., Yu, W., & Ji, Z. (2013). An exploration of Chinese EFL students' emotional intelligence and foreign language anxiety. *The Modern Language Journal*, *97*(4), 917-922. https://europepmc.org/article/med/31736840

Tomic, W., Evers, G., & Brouwers, A. (2004). Existential fulfillment and teacher burnout. *European Psychotherapy*, *5*(1), 65-73. https://research.ou.nl/en/publications/existential-fulfillment-and-teacher-burnout

Triandis, H. (2005). Cultural intelligence in organizations. *Group & Organization Management*, *31*(1), 20-26. https://doi.org/10.1177/1059601105275253

Van Dyne, L., Ang, S., & Koh, C. (2008). Development and validation of the CQS. In S. Ang & L. Van Dyne. (Eds.), *Handbook of cultural intelligence*: *Theory, measurements, and applications* (pp. 16-40). M. E. Sharpe. https://www.taylorfrancis.com/books/9781315703855/chapters/10.4324/9781315703855-10

Van Dyne, L., Ang, S., & Koh, C. (2009). Cultural intelligence: Measurement and scale development. In M. A. Moodian (Ed) *Contemporary leadership and Intercultural competence: Exploring the cross-cultural dynamics within organizations* (pp. 233-254). Sage. https://doi.org/10.4135/9781452274942.n18

Van Dyne, L., Ang, S., & Livermore, D. (2010). Cultural intelligence: A pathway for leading in a rapidly globalizing world. In K. Hannum, B. B. McFeeters, & L. Booysen (Eds.), *Leading across differences* (pp. 131-138). Pfeiffer. https://doi.org/10.1.1.582.29608

Van Dyne, L., Ang, S., Ng, K. Y., Rockstuhl, T., Tan, M. L., & Koh, C. (2012). Sub-dimensions of the four factor model of cultural intelligence: Expanding the conceptualization and measurement of cultural intelligence. *Social and Personality Psychology Compass*, *6*(4), 295-313. https://doi.org/10.1111/j.1751-9004.2012.00429.x

Van Dyne, L., Ang, S., & Nielsen, T. (2007). Cultural intelligence. *International Encyclopedia of Organization Studies*, *1*, 345-350. https://culturalq.com/wp-content/uploads/03/MOR-2007-Ang_Van-Dyne-etc.pdf

Wallace, J. E., & Brinkerhoff, M. B. (1991). The measurement of burnout revisited. *Journal Social Service Research*, *14*(5), 85-11. https://doi.org/10.1300/J079v14n01_05

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Investigating Iranian EFL Learners' Negotiation Types in Writing with Different Scaffolding Patterns

Research Article
pp. 95-110

Aysheh Mohammadzadeh[1]
Touran Ahour*[2]
Mahnaz Saeidi[3]

Received: 2020/02/17 Accepted: 2020/09/26

Abstract
Learner-centered approaches in second language acquisition and process approach in writing pedagogy has stimulated quite a number of researchers to focus on learners' voices in collaboration passing through multiple drafts and revisions. This study based on the concept of scaffolding learning in Vygotsky's sociocultural theory and process writing approach to second language writing investigated Iranian EFL learners' negotiation types in small groups of different scaffolding patterns; symmetrical and asymmetrical. To this end, 15 students at High and Low Intermediate Proficiency levels were assigned into three groups in different scaffolding patterns; one asymmetrical group with two High Intermediate - three Low Intermediate learners (H-L), two symmetrical groups with five High Intermediate learners (H-H), another with five Low Intermediate learners (L-L). Small group interactions were observed and recorded. Transcriptions were analyzed to identify negotiation types in terms of language functions among different groups. Therefore, two main categories, Responding and Requesting and their subcategories were found. The subcategories of "agreeing", "explaining", "giving opinions", "instructing", "restating" and" suggesting" were related to the first main category; "comprehension checking", "eliciting opinions" and "questioning "were related to the second main category. According to Chi-square test results, negotiation types were significantly related to the scaffolding pattern as the asymmetrical (H-L) group was superior in the number of language functions used over the symmetrical (L-L, H-H) groups. Nevertheless, members in all groups enjoyed high equality and mutuality in interaction. The findings suggest teachers raise learners' awareness of the diverse strengths and abilities that different scaffolding patterns give them.

keywords: asymmetrical scaffolding, collaborative writing, EFL learners, negotiation types, symmetrical scaffolding

* Corresponding Author
[1] PhD Candidate, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran. Mohammadzadeh79@gonbadiau.ac.ir
[2] Assistant Professor, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran. ahour@iaut.ac.ir; touranahour@yahoo.com
[3] Associate Professor, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran. m_saeidi@iaut.ac.ir

DOI: 10.22051/LGHOR.2020.30357.1259

Introduction

During the last three decades, the Process Writing Approach (PWA) has gained very much attention in the second language writing. Within the writing process, collaborative writing (CW) in which learners work together in order to accomplish a writing task, has roots in Vygotsky's sociocultural theory and the concept of scaffolding. CW focuses on interaction by activating the social resources of the learners (Storch, 2005) and emphasizes students' negotiation of meaning which leads to their accountability for their own learning and decision-making power in the class (Breen & Littlejohn, 2000). In other words, it asserts that students are not passive, but rather they learn through active, dialogic interactions with their teachers, peers, and the context (Lantolf, 2000). Accordingly, researchers have been impressed by interaction in small writing group tasks in the last few decades (Storch, 2002).

Nevertheless, reviewing the literature indicates that the use of CW in writing classes is not emphasized in the Iranian English as a Foreign Language (EFL) context (Biria & Jafari, 2013; Jafari & Nejad Ansari, 2012), and the number of studies specifically regarding the investigation of negotiation types in the interaction of small groups of different scaffolding patterns is scant. Some studies in the context of Iran (Biria & Jafari, 2013; Jafari & Nejad Ansari, 2012; Khodabakhshzadeh & Samadi, 2018; Soleimani et al., 2015) compared collaborative work with individuals or paired work but they did not investigate interactive patterns and negotiation types altogether. However, the nature of interaction in writing tasks while learners working within small groups, deserves further investigation.

In this regard, the present study based on Vygotsky's sociocultural learning theory aimed at uncovering what happens in small interactive writing groups regarding the negotiation types in terms of language functions. Moreover, it investigated EFL students' interaction during a writing task in small groups of different scaffolding patterns namely symmetrical and asymmetrical.

Literature Review

The Process Writing Approach

The Process Writing Approach (PWA) used in classrooms today is a teaching approach which originates from Piaget's constructivist theory and also from Vygotsky's socio-cognitive theory which focuses on 'zone of proximal development' (ZPD) and 'more knowledgeable other' (MKO) concepts. Scaffolding, collaborative learning and apprenticeship are more concepts from Vygotskian theory. Social cognitive theory also sheds light on social, affective and motivational components of learning (McCutchen et al., 2008).

The effectiveness of PWA is attributed to the cognitive activities when students plan, draft and revise (Rijlaarsdam & Van den Berg, 2006). Graham and Sandmel (2011) hypothesize that its effectiveness is also due to using some components to improve writing such as mini-lessons and writing conferences, and methods to increase motivation such as collaboration and positive learning environment. In second language (L2) classrooms where collaborative learning is emphasized, learners are involved at different proficiency levels, so exploring the effects of proficiency differences on interaction is prevalent.

Sociocultural Theory

Vygotsky's (1978) sociocultural theory (SCT) places social context at the

heart of the learning and communication process. According to his theory, students' ability to learn and write does not occur only through their own personal and individual cognitive activities. It is also influenced and contributed both - consciously and unconsciously by individuals' social and cultural context that surrounds them in the same way. In other words, for an individual to become a proficient learner, the mastery of the language is achieved through people's participation to negotiate meaning (Fahim & Haghani, 2012). Moreover, the theory emphasizes the interconnections among teachers, learners and tasks, and encourages interactions among individuals (Fung, 2006). Ellis (2000) states that the interaction helps learners to scaffold the new tasks while they are in the process of learning.

Conceptual and cultural learning occur through dialogue in Vygotsky's (1978) zone of proximal development (ZPD), that is, the difference between the actual developmental level and the potential development level under the guidance or in collaboration with a more abled peer. It is important to note that learning in a ZPD may be effectively scaffolded by either teachers or fellow learners. However, the concept of ZPD has been the subject of many studies. Ohta (2001), for instance redefined it in second language learning as "the distance between the actual developmental level as determined by individual linguistic production, and the level of potential development as determined through language produced collaboratively with a teacher or peer" (p. 9).

Scaffolding and its Different Patterns. Scaffolding is typically associated with the sociocultural theory of Vygotsky (1978). It is an instructional technique whereby the teachers or more proficient participants model the appropriate learning strategy or task and gradually give the responsibility to the less proficient students. As the interaction continues, the interpretations are refined and reconciled dialectically in the process of mutual learning.

Scaffolding is a general concept with different divisions. Two types of scaffolding include the symmetrical and asymmetrical forms. "Symmetrical scaffolding rests on the fact that learners discover new knowledge through cooperation and interaction", (Baleghizadeh et al., 2010, p. 105). In other words, students with the same ZPDs cooperate with each other (Baleghizadeh, et al., 2010).

In the asymmetrical scaffolding, the learners with different ZPDs work with each other. In fact, asymmetrical scaffolding is a typical kind of scaffolding in which there is always a learner who is more knowledgeable than others (Baleghizadeh et al., 2010). As a result, the concepts of scaffolding and ZPD principally encompass the provision of asymmetrical scaffolding in teaching and learning.

Negotiation in Collaborative Writing

Many researchers (e.g., Biria & Jafari, 2013; Fahim & Haghani, 2012; Farrah, 2012; Storch, 2005) have indicated that in a collaborative writing (CW), writers are active in decision making about the language they want to state their ideas and making the structure to state those ideas together. Wells (2000), states that "knowledge is created and re-created in the discourse between people doing things together" (p. 71). Thus, knowledge building is the result of collaborative work through discourse and "the constructive and creative effort involved in saying and in responding to what was said" (Wells, 2000, p. 74).

According to Ellis (2000), when L2 learners have communicative problems and at the same time have the opportunity to negotiate with others, they can acquire

language better. Negotiation is therefore necessary to make the input comprehensible. It can enhance mutual understanding and contribute to the development of learner autonomy (Breen & Littlejohn, 2000). Learners in this way, have the opportunity to consider their own learning process and share resources. According to Yu (2008), "the notion of negotiation is generally defined as the discussion to reach agreement" (p. 48). Negotiation has an important role in the interaction of the classroom because the learners find more opportunities to negotiate their problems (Yu, 2008).

Studies of adult language learners (e.g., Shortreed, 1993; Yule & MacDonald, 1990) have indicated that the amount of negotiating for meaning is closely related to the proficiency level. That is, Non-Native Speaker_ Native Speaker (NNS–NS) pairs negotiate for meaning more than their Native Speaker_ Native Speaker (NS–NS) counterparts in order to resolve misunderstandings between partners. On the other hand, other studies (e.g., Foster & Ohta, 2005; Ohta, 2001) claim when learners support each other, share meanings, and modify their own and each other's utterances while monitoring, more ZPDs are created as learners rely on one another to proceed regardless of their language proficiency. So, grouping learners in different patterns of learning for example scaffolding patterns or according to their language proficiency to examine how negotiation proceeds, seems necessary in this regard.

Empirical Studies

Although CW is widely studied in EFL writing instruction, little is known about the nature of the interaction and negotiation types between peers in small different scaffolding groups. A few studies in the literature have inquired negotiation types in pairs not in groups. Mendoca and Johnson's (1994) seminal work is almost the first study that investigated peer review negotiations. Twelve ESL learners who were advanced international graduate students with different educational fields took part in the study. Audio-taped peer review transcripts and also first and revised drafts analyses, and then post-interviews were gathered**.** The study found that students used peers' comments in revising the essays and certain types of negotiations like questioning, explaining, suggesting, restating, and correcting grammar mistakes occurred frequently. Post interviews revealed that generally they found peer reviews useful.

Storch's (2002) influential study investigated interaction patterns in a context of adult second language learning. The data were obtained from 10 pairs in three tasks; composing, editing and text reconstructing. Pair talks were audio-recorded and salient traits were analyzed. Four patterns were emerged describing the role relationships; collaborative, dominant/ dominant, dominant/passive, and expert/ novice. According to the results, collaborative pattern predominated and was stable regardless of the task and time. In addition, the transfer of knowledge was more evident in collaborative and the expert/novice dyads than in dominant/dominant and dominant/passive dyads. Storch explained these results referring to the cognitive development theory of Vygotsky as members co-constructed, appropriated, and internalized the knowledge.

Watanabe (2008) asserted the importance of interaction, in contrast to language proficiency, in the writing performance of learners. Students participated in a writing task with higher and lower proficiency levels. The transcribed pair talks were analyzed regarding words, language episodes, and pair interaction patterns. He

demonstrated higher- and lower-proficiency pairs both enjoyed chances of learning in a collaborative pattern, and shared reciprocal ideas and contributed in writing equally.

Memari Hanjani and Li (2014) explored learners' interaction in a collaborative revision task and its impact on the writing performance. Some pairs of L2 students who enrolled in an essay-writing course participated in the study at a university in Iran. Each pair took part in a collaborative revision session and revised argumentative texts jointly. Students applied different functions in their negotiations. Their revision interactions included evaluative (scaffolding and non-scaffolding), social (on-task and off-task), and procedural negotiations. It was revealed that the joint revision task was beneficial for both partners.

Li and Kim (2016) investigated the interactions of two ESL groups in two collaborative writing tasks working in a Wiki space. They examined language functions which learners employed during task negotiation, writing change functions and scaffolding strategies. The two groups' wiki activities were recorded as main data sources. The results drawn on sociocultural theory showed that learners utilized the target language as a mediating tool in interacting with group members to perform writing tasks and to negotiate social relationships. Exploring writing change functions showed an ongoing joint writing process at both writing and revising stages. Unlike Storch (2002) who discovered relatively stable interaction patterns that pairs revealed, Li and Kim (2016) indicated the changing of interaction patterns and the fluidity of scaffolding.

Li and Zhu (2017) examined the connections between writing products and interaction patterns in the wiki writing task environment. According to the results, the group with the collective pattern, produced a higher writing quality, especially in the rhetorical structure and coherence. Next group with high quality showed an expert/novice pattern. Groups which showed a dominant/defensive and a cooperating-in-parallel pattern produced research proposals of relatively low quality. Interactions in Wiki and writing products were linked to the concept of scaffolding.

Although the findings of these studies support negotiation in writing, research is still needed to consider the nature of negotiation in different patterns of learning; for instance, symmetrical and asymmetrical scaffolding patterns. Reviewing the previous literature revealed that mostly peer interaction in revising the writing tasks has been explored. Few studies have probed negotiation in small groups with regard to different patterns of scaffolding. Therefore, the present endeavor attempted to bridge this gap and probe what types of negotiation EFL students engage in while working in small groups of different scaffolding patterns; that is, symmetrical with (High-High) H-H, (Low-Low) L-L groups and asymmetrical with (High-Low) H-L group.

According to Ohta (2001), learners involving in collaborative dialogues, may pool their strengths and weaknesses and co-construct more knowledge as a group, regardless of their high differences or degrees of homogeneity in proficiency levels. Therefore, scaffolding patterns were scrutinized in this study to see the nature of collaborative orientation of groups or how interaction is shaped among learners. In this regard, the following research questions were posed:

RQ1: What types of negotiation are used in small writing groups with different patterns of scaffolding?

RQ2: Is negotiation type significantly related to the scaffolding pattern?

RQ3: What are patterns of interaction that occur in each scaffolding group?

Method

Participants

The present study was drawn from a larger study, which probed the effects of symmetrical and asymmetrical scaffolding on L2 writing fluency, complexity and accuracy. In order to have groups of different scaffolding patterns, a pool of 117 intermediate female EFL learners from seven intact classes at a language institute in Gonbad Kavous, Iran took part in the study. According to the results of Preliminary English Test (PET), 90 students whose scores fell one standard deviation below (low proficiency = 45) and above (high proficiency = 45) the mean were chosen as the participants of the study. Next, they were randomly assigned to three groups (30 in each group), namely; one asymmetrical group with 15 high and 15 low level students, one symmetrical group with 30 high level students, and another symmetrical group with 30 low level students. Then, the members of three participant groups were divided into the small groups. For the purpose of this study one group from each pattern was selected. So, generally 15 participants worked in three groups (5 students in each group) with three different scaffolding patterns.

The participants were within the age range of 18-22 and had been studying English at the institute at least for two years. They had started learning English as a compulsory subject since grade seven at schools. They were all Persian native speakers and had no formal, systematic previous exposure to any writing courses or collaborative writing activities. With respect to the fact that the study was conducted in a private language institute, the socioeconomic status of the participants was deemed to be middle to upper-middle.

Instruments and Materials

Preliminary English Test (PET). PET tests the test takers' writing, reading, listening and speaking skills. The reading and writing sections were taken together in 90 minutes. The listening part lasted for 30 minutes and the interview stage was conducted in 10 minutes. The maximum score on this test is 170.

Observation. In order to probe the types of negotiation used in the interaction of the small groups in three scaffolding patterns, observation was employed. That is, the researchers placed one digital audio recorder next to each group to audio-record one session of each small group (as a focus group) in the symmetrical and asymmetrical scaffolding patterns.

Procedures

At the outset of the study, the purpose and the learners' right to withdraw from the study at any time was explained to the participants. In all groups, participants were informed of the collaboration rules, for instance, having complementary roles, planning, generating, suggesting alternative ideas, and listening to each other, etc. They were also instructed with some rules of composing cohesive and unified one-paragraph essays. The teacher asked them to write a descriptive composition on a given topic by discussing with each other. Each session, they were asked to write a maximum of 200-250-word one-paragraph essay. In other words, they were asked to interact with each other and collaboratively write the compositions.

Students' writing process took 30 minutes in each session. In the asymmetrical class different level learners in small groups worked together to achieve the purpose of the group work, and in the symmetrical classes, either low

level learners or high level learners worked together for this reason.

Learners underwent three phases while composing; planning, composing, and revising. In the planning phase which took longer than the other two phases, they brainstormed on the topic before starting their writings. Participants shared their ideas by discussing the content and organization of the writings in this phase. In the revising phase, they did not spend much time as they corrected the errors and provided feedback in the composing phase. Since the study aimed at capturing the types of negotiation taking place in each group, an audio recording was carried out while learners worked together passing the different phases.

For the purpose of this study the fourth session was audio recorded. Because the learners were experiencing writing in a group for the first time, the first three sessions served as just warm up. Due to the large number of groups working at the same time, the audio recording was based upon one focus group in each scaffolding pattern to find negotiation types between the members.

Students were to select among these topics in the fourth session: *"what makes a film great?", "description of a desert", "life in crowded cities",* and *"my favorite restaurant"*. The audio recording was made by the researchers during the time that participants were working to complete the writing tasks.

Design and Data Analysis

Regarding the first research question, the interactions of the group members in each scaffolding pattern were analyzed qualitatively. So the focus group audio-recordings were listened to and transcribed using the standard orthography. Their contents were analyzed inductively for language functions that were used in the interaction between the members of the small groups for coding instances of negotiation. Descriptive categories were concluded through a grounded approach in which subcategories are examined, compared and connected to each other to identify core categories (Strauss & Corbin, 1998). The inter-rater reliability agreement came 85% by an outside researcher as an acceptable level of coding reliability. Therefore, emerging categories were compared in three groups to find the similarities and differences in the negotiation of three scaffolding patterns (H-L, H-H, and L-L).

As for the second research question, descriptive statistics including frequency of the negotiation types in each group were determined and then they were compared quantitatively through Chi-Square to find whether negotiation types were significantly related to the scaffolding pattern or not. To answer the third research question, patterns of interaction were analyzed based on two indexes of peer engagement; that is, equality and mutuality (Storch, 2002). Equality means all members taking control rather than submitting to a one-way direction from others in contributing the task; and mutuality means all members enjoying the same level of engagement in the contribution. The aim was to discover the nature of relationships learners had in different scaffolding patterns.

Results

RQ1: What Types of Negotiations Are Used in the Three Scaffolding Groups?

Small groups' interactions were examined through analyzing group members' negotiation regarding language functions performed during the writing task. In this way, two main categories and their subcategories were found. Table 1 presents the definitions for these categories and subcategories found in this regard.

Table 1
Taxonomy of Language Functions Found in the Groups' Interactions

Language Functions	Definitions
Agreeing	Expressing agreement with other members' ideas
Explaining	Explaining the meaning of a term or idea that is not clear to each other
Giving opinions	Members giving ideas or comments
Instructing	Teaching, for example, grammatical structures or writing the essay
Restating	Rephrasing what has been written or said to show understanding or rereading sections
Suggesting	Suggesting or recommending other ways to change the words, content, or organization
Comprehension checking	Checking for understanding the meaning of a term or idea or what has been said
Eliciting opinions	Members draw ideas or comments
Questioning	Asking questions where something is unclear
Main categories	
Responding	Giving answers or reacting to other members' requests
Requesting	Making requests where something is unclear

According to Table 1, among language functions, "agreeing", "explaining", "giving opinions", "instructing", "restating" and" suggesting" were related to the first main category of Responding. So, "comprehension checking", "eliciting opinions" and "questioning "were related to the second main category, that is, Requesting.

RQ2: Is Negotiation Type Significantly Related to the Scaffolding Pattern?

In order to answer the research question two, a Chi-square test for independence was applied. Table 2 shows the results of the Chi-square test, which include the frequency of the nine language functions for the three scaffolding groups.

Table 2
Frequency of Language Functions among Groups

Language Functions	Group			Total	Chi-Square	df	Sig
	H-L	H-H	L-L				
Agreeing	11	10	4	25	86.608	16	.000
Explaining	51	19	21	91			
Giving ideas	31	36	15	82			
Instructing	49	25	20	94			
Restating	53	27	18	98			
Suggesting	50	49	22	121			
Comprehension checking	44	20	19	83			
Eliciting opinion	55	14	10	79			
Questioning	20	25	48	93			
Total	364	225	177	766			

As the results of Chi-square test for a (3 scaffolding patterns and 9 language functions) contingency table (Table 2) shows, there was a significant relationship between the scaffolding patterns and negotiation types, $\chi^2 = 86.61$, df =

16, p = .000.

Based on the results, compared to the H-H and L-L groups, the H-L group has the highest frequency in all functions except for the functions of "giving ideas" and "questioning" that are frequently used functions in the H-H and L-L groups, respectively. As six and three language functions are, respectively, related to the categories of Responding and Requesting, another Chi-square test was carried out to find the relationship between the main categories and scaffolding patterns. Table 3 indicates the results.

Table 3
Frequency of the two Categories among Groups

Language Functions	Group			Total	Chi-Square	df	Sig
	H-H	H-L	L-L				
Responding	166	245	100	511			
Requesting	59	119	77	255	13.433	2	.001
Total	225	364	177	766			

The results in Table 3 demonstrate a significant relationship between the two main categories and the three scaffolding patterns (2×3 contingency table), χ^2 = 13.43, df = 2, p = .001. As the results reflect, while the category of Responding has the highest frequencies in all three scaffolding groups, the H-L group has the highest frequency of using the functions of Responding and Requesting compared to the H-H and L-L groups. In this regard, the second research question is answered positively.

RQ3: What Are Patterns of Interaction That Occur in Each Scaffolding Group?

To answer the third research question regarding the interaction patterns in each group, Storch's (2002) model was adapted that used indexes of "equality" (i.e., the degree of contribution and control over the direction of writing) and "mutuality" (i.e., the extent of engagement with each other's contribution). In her influential work on the nature of peer interaction in collaborative writing tasks, Storch (2002) introduced four types of interaction patterns in ESL pair writing processes: collaborative, dominant/dominant, dominant/passive, and expert/novice.

In this regard, frequency and language function types that members performed in small groups were examined (Li & Kim, 2016; Li & Zhu, 2017). Instances of contributions from all members of the group in terms of language functions suggested high equality. That is, equality was defined according to the degree of contribution of all members while working on the task. Whether the group enjoys balanced contribution or only some members take the control of the flow of negotiation by dictating their ideas was the concern of equality. Mutuality was examined by frequency of language functions such as giving opinions, suggesting, instructing, and agreeing. The goal was to discern if all members engage with each other to proceed the task.

Group 1 (H-L). Comparing the higher number of functions, negotiation seems to proceed better in this asymmetrical group than the other two symmetrical groups. This can be ascribed to the high ability members' understanding that their low level peers had problems in group interactions or task completions and therefore attempted to assure whether they could reach their level by consistently checking

their comprehension. In the excerpts 1 and 2 in the planning phase, there are samples of "instructing", "explaining" and "comprehension checking" that suggest high ability members' effort to actively participate in the task along with all low ability members. Generally, "eliciting opinions", "restating", "explaining", suggesting", "instructing", and "comprehension checking" were used more in this group in comparison to other two symmetrical groups.

The interaction pattern in this group was tutor/tutee or expert/novice, as the number of language functions related to the category of Responding such as "eliciting opinions", and language functions for the category of Requesting like "restating", "explaining" and "suggesting" exceeded other language functions. However, low ability members also actively collaborated in the task as they felt free in "questioning" as a sign of equality in the group. That is, contribution to the task came from all directions not from only high ability members. As a result, the group enjoyed high equality as all members contributed jointly to the task and none took the control. In addition, they used language function of "giving ideas", as well as samples of "agreeing". Therefore, while high ability members expected to take the control of the task, low ability members were not passive (high mutuality). According to Storch (2002) the expert/novice interaction pattern is collaborative in nature. Therefore, it can be claimed that the collaborative pattern was evident in this group.

In the following excerpts A, B, and E are high, and C and D are low ability members. The topic selected by this group was: *"life in crowded cities.*"

Excerpt 1:
A: Well, I'm not sure too much emotional moment make the film boring (*giving an opinion*), maybe if it ends too quickly I get bored (*explaining*)…which do you think is important? (*comprehension check*)
B: I think….I agree (*agreeing*) because when you don't understand the end of a film you become bored (*restating*) and think that waste your money (*giving an idea*).
A: Right? (*Comprehension check*) Once it happened to me (acknowledging) …I don't like going to cinemas (*giving an opinion*)…I usually buy CDs and watch in home (*explaining*).
C: What was that boring film for you? (*questioning*)
D: Ayneh Baghal [Side-view Mirror: the name of the film] (*explaining*)
A: No need to mention the names of the movies in the essay instead we should mention the characteristics. (*instructing*)

Excerpt 2:
A: Let's write about actors and actresses (*suggesting*)
C: They act should look real. (*giving an opinion*)
B: Right (*agreeing*)
D: Yes (*agreeing*)
E: Directors are important too, someone like Asghar Farhadi. Everybody likes his movies. (*giving an opinion*)

Group 2 (H-H): Analyzing some selected excerpts, it is assumed that some H-H peers did not dominate others through dictating their own opinions. All members were at ease in "giving opinions" and "suggesting" rather than "explaining" or "eliciting opinions", since they had a good command over the topic and interactions. Due to the equal roles members had regarding each other, directions

equally came from all members. Since all members enjoyed balanced contributions to the task the group enjoyed high equality. While some members presented opinions and gave new suggestions, others agreed and restated them to outline related ideas jointly. Therefore, the group enjoyed high mutuality as members engaged in the task using different language functions such as "suggesting" and "giving ideas" to accomplish it collaboratively.

Hence, according to Storch (2002), the interaction pattern in this group was collaborative. In the following episodes of excerpts; for instance, members all contribute to the task in the planning phase. There are agreements and contributions to the suggestions and restating them by other members. So it is claimed that the group was high in equality and mutuality. The topic chosen was *"My favorite restaurant"*:

Excerpt 3:
D: We can start with different kinds of restaurants. (*suggesting*).
A: Restaurants which serve traditional food are healthier (*giving an opinion*).
B: Yes I see (*agreeing*), I love fast food too (*giving an opinion*)
C: You can have traditional food at home and fast food out (*suggesting*).
D: what about the location? (*eliciting an opinion*)
E: It's very important, clean and modern, city center or out in the country? (*suggesting)*
A: Ok, so let's write about first different kinds, then... (*restating*)
B: it's very important the restaurant be clean…clean kitchen, clean plates, and clean environment. (*restating*)
B: I agree with you there. (*agreeing*)
C: Don't forget the variety in food. (*suggesting*)
D: You can find variety in just larger cities like Tehran. (*giving an opinion*)
E: Yes, nothing about Chinese or Italian foods here in our city (*explaining*). We can end the essay with this point. (*suggesting*)

Excerpt 4:
C: and bad points? (*suggesting*)
D: Why do you say that you like it? (*questioning*)
A: My favorite restaurant cooks not just fast food, but also some traditional food like GhormehSabzi. (*explaining*)
C: Does it accept phone orders? (*questioning*)
A: Yes, but sometimes it gets late in bringing the food and the food becomes cold! (*explaining*)
D: I think if they can't bring food on time, they shouldn't accept home orders (*giving an opinion*).
C: So we can write it as a negative point. (*restating)*

Group 3 (L-L). Although lower number of language functions was found in this group, none was passive in terms of equality among members. All participants contributed to accomplish the task (high equality). As samples of language functions used in this group reveal in the following excerpts, in terms of mutuality, members' using "suggesting" and "instructing" demonstrated a good degree of mutuality. Hence, the dominant interaction pattern was collaborative in this group, too. The following excerpts in the planning phase show some samples of negotiation in this group. The topic was *"life in crowded cities*".

Excerpt 5:
A: What about stick in traffic…Is it correct? (*questioning*)
B: Yes, it's right (*agreeing*). I heard it before, sticking in traffic (*restating*)
A: I think it's not enough to just say you don't like to live in crowded cities (*giving an opinion*)….you should give some explanation and examples and compare the disadvantages and advantages of life in cities to country side (*explaining*)…for example, sometimes I find the heavy traffic too much and want to free up my mind from all the routine works (*suggesting*).
B: Yes, like crowded streets, big supermarkets (*explaining*)…what do you say if you want to tell very heavy traffic? (*questioning*)
A: I heard bumper to bumper (*suggesting*). When you can't move your car and you wait for hours on the road then you can say traffic jam and bumper to bumper (*explaining*).

Excerpt 6:
A: for example, like Tehran…it's a crowded city and always full of people. If I live there I wish to escape and go to some green places…maybe parks.
B: Yes, a lot of trees in the park help people in crowded cities enjoy their life (*restating*).
C: Or you can go to different national parks outside the city… (*suggesting*).
D: Yes, it's great.

Excerpt 7:
A: I note all negative points about large cities.
B: Maybe…we could also talk about positive points and in this way we make the content more understandable and others can compare the advantages and disadvantages for themselves…also if we only talk about bad points, it looks like we hate cities, but I think I like some things about them (*suggestion*).
D: Like what? (*questioning*)
C: For example, we can say a lot of schools, shopping malls, entertainment, hospital and etc. (*suggesting*).

Discussion

This study attempted to document small groups' negotiation types in the interaction of different scaffolding patterns while writing in the second language. Unlike most previous studies (e.g., Memari Hanjani & Li, 2014; Mendoca & Johnson, 1994; Storch, 2002) that investigated patterns of interaction in pairs in the revision phase, the present study explored interaction patterns in the negotiation of learners working in small groups accomplishing the writing task in different phases mostly in the planning phase.

Analyzing members' negotiation types in terms of language functions, two main categories (i.e., Responding and Requesting) and their subcategories (i.e., agreeing, explaining, giving opinions, instructing, restating, suggesting, comprehension checking, eliciting opinions, and questioning) were found, which were mostly in common with Li and Kim (2016) and Mendoca and Johnson (1994). In addition, some of the language functions found in this study such as "instructing", "restating", and "comprehension checking" overlapped Memari Hanjani and Li's (2014) scaffolding and non-scaffolding language function respectively, which they called them as evaluative negotiations.

Results of Chi-square demonstrated that negotiation types were significantly related to the scaffolding pattern. The asymmetrical (H-L) group was superior in the number of negotiation types used over other symmetrical (L-L, H-H) groups. This suggests negotiation proceeds more successfully in the asymmetrical type of scaffolding pattern in which learners who are less able benefit from the help of more competent members in a supportive group environment. This can be explained according to Vygotsky's concept of scaffolding which asserts the social aspect of learning in a joint collaboration between more knowledgeable and less knowledgeable individuals. In this group, "explaining", "restating", "suggesting" and "eliciting opinion" found to be the most frequent negotiated language functions, suggesting that from a social point of view, negotiation needs to proceed by the smooth flow of exchanges among the members of the group in which more competent members try to scaffold less competent ones. For instance, less able members' requests for more explanations suggest that they looked to more able ones to get more feedback.

"Giving ideas" and "suggesting" occurred more frequently in the H-H group with the symmetrical pattern. This is in line with Shortreed (1993) and Yule & MacDonald (1990) that found more proficient second language learners' pairs display features similar to the characteristics of Native Speaker- Native Speaker (NS- NS) dyads than Native Speaker- non-Native Speaker (NS- NNS) dyads probably due to their confidence over the topic. That is, NNS–NS dyads negotiate for meaning more than their NS–NS counterparts due to misunderstandings between language partners. The frequency of negotiation types also proved this finding in that they were less in number in comparison to the H-L group. Accordingly, it is suggested that grouping students in asymmetrical patterns creates a high supportive context. In contrast, in the L-L group, the total number of the negotiation types was the least. The most frequent language function in this group was "questioning" which suggests that too many requests for clarification may create uneasy social relationships.

Nevertheless, in the present study, the analysis of interaction patterns (as evident in sample excerpts) supported that students in all three groups with different scaffolding patterns were capable of providing assistance and feedback in writing as they enjoyed high mutuality and equality. Different scaffolding groups of the study, fit within Storch's (2002) collaborative pattern of interaction: a collaborative pattern (collaborative and expert/novice) against a non-collaborative pattern (dominant/dominant and dominant/passive).

Although the H-H and H-L groups met this general expectation, the unexpected finding was revealed to be the L-L members' exemplification of the collaborative pattern of interaction, too. This indicates that social mediation comes not only from experts such as teachers and more capable peers but also even from less proficient peers. This is in line with Ohta (2001) who stated that a learner is not universally less or more able than his/her peer and each learner shows a collection of weaknesses and strengths that might be complementary.

Moreover, in contrast to previous studies (Memari Hanjani & Li, 2014; Mendoca & Johnson, 1994; Storch, 2002) which found peers' concern over forms (e.g., structure and wording) in the revision phase, the findings of this study are in line with Li and Kim (2016). They claimed when academic writing tasks focus on the application of emerging genre knowledge, language forms receive much less attention. Although the participants in this study were novice writers, few samples

related to language forms were found. Instead, members used different language functions related to the content topic at least in the planning phase, as analyzed in the episodes of different group members' utterances.

Conclusion

This study was an attempt to document second language learners' negotiation types in writing while interacting in small groups with different scaffolding patterns. Negotiation types, analyzed in terms of language functions, were significantly related to the scaffolding pattern, so that the asymmetrical (H-L) group was superior in the number of negotiation types used such as "eliciting opinions", "restating", "explaining", suggesting", "instructing", and "comprehension checking" over other symmetrical groups (L-L, H-H). This finding suggested the successful proceeding of negotiation in the asymmetrical type of scaffolding pattern in which learners who are less able benefit from the help of more able members in a supportive group context. However, the collaborative nature of interaction in all three groups proved that social mediation can come even from less proficient peers.

The findings of this study would be of great benefit to practitioners and theoreticians in the field of language teaching in general and teaching writing in particular. EFL teachers can encourage the learners' autonomy by providing insights on the scaffolding and collaboration in EFL settings. They can provide opportunities for learners to interact and work collaboratively in small groups of asymmetrical and symmetrical patterns. Thus, it is thought that investigating CW would help language teachers to be aware of the effectiveness of scaffolding in an EFL context. The findings also suggest teachers to create an atmosphere of interdependence among the EFL learners by encouraging them to support one another spontaneously with their diverse strengths and abilities. Students also benefit from the results of this study as CW in different scaffolding patterns gives them a sense of accountability for their group members' engagement.

Due to some limitations, the participants of this study were only among female EFL learners. Exploring the role of gender in group interaction and its possible association with the nature of negotiation in scaffolding learning patterns also looks appealing, as it may provide opportunities for a better understanding. Especially in an EFL context like Iran where education is affected by cultural norms and religious beliefs of the society, probing the effect of gender can be quite required and open new insights to the field.

References

Baleghizadeh, S., TimcheMemar, H., & TimcheMemar, A. (2010). The effect of symmetrical versus asymmetrical scaffolding on English reading comprehension of EFL learners. *Studies in Literature and Language, 1*(7), 104-111. www.cscanada.org

Biria, R., & Jafari, S. (2013). The impact of collaborative writing on the writing fluency of Iranian EFL learners. *Journal of Language Teaching and Research, 4*(1), 164-175. https://doi.org/10.4304/jltr.4.1.164-175

Breen, M. P., & Littlejohn, A. (2000). *Classroom decision making*. Cambridge University Press.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4*(3), 193-220. https://doi.org/10.1177%2F136216880000400302

Fahim, M., & Haghani, M. (2012). Sociocultural perspectives on foreign language learning. *Journal of Language Teaching and Research, 3*(4), 693-699. https://doi.org/10.4304/jltr.3.4.693-69

Farrah, M. (2012). The impact of peer feedback on improving the writing skills among Hebron university students. *An-Najah University Journal for Research-B (Humanities), 26*(1), 179-210. http://scholar.najah.edu/publication/journal-article/impact-peer-feedback-improving-writing-skills-among-hebron-university

Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied Linguistics*, *26*(3), 402-430.

Fung, M. Y. (2006). *The nature and dynamics of collaborative writing in Malaysian tertiary ESL Setting* (Unpublished doctoral dissertation). Massay University, Palmerston, North New Zealand.

Graham, S., & Sandmel, K. (2011). The process writing approach: A meta-analysis. *Journal of Educational Research*, *104*(6), 396-407. https://doi.org/: 10.1080/00220671.2010.488703

Jafari, N., & Nejad Ansari, D. (2012). The effect of collaboration on Iranian EFL learners' writing accuracy. *International Education Studies, 5*(2), 125-131. https://doi.org/10.5539/ies.v5n2p125

Khodabakhshzadeh, H., & Samadi, F. (2018). The effect of collaborative writing on Iranian EFL learners' task achievement in writing and their perception. *International Journal of Applied Linguistics & English Literature*, *7*(1), 113-119. http://dx.doi.org/10.7575/aiac.ijalel.v.7n.1p.113

Lantolf, J. (2000). *Sociocultural theory and second language learning*. Oxford University Press.

Li, M., & Kim, D. (2016). One wiki, two groups: Dynamic interactions across ESL collaborative writing tasks. *Journal of Second Language Writing, 31*(1), 25-42. https://doi.org/10.1080/00220671.2010.488703

Li, M., & Zhu, W. (2017). Good or bad collaborative wiki writing: Exploring links between group interactions and writing products. *Journal of Second Language Writing, 35*(2), 38-53. https://doi.org/10.1016/j.jslw.2017.01.003

McCutchen, D., Teske, P., & Bankston, C. (2008). Writing and cognition: Implications of the cognitive architecture for learning to write and writing to learn. In C. Bazerman (Ed.), *Handbook of writing research* (pp. 451-470). Lawrence Erlbaum.

Memari Hanjani, A., & Li, L. (2014). Exploring L2 writers' collaborative revision interactions and their writing performance. *System*, *44*(4),101-114. https://doi.org/10.1016/j.system.2014.03.004

Mendoca, C., & Johnson, K. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly, 28*(4), 745-769. https://doi.org/10.2307/3587558

Ohta, A. S. (2001). *Second language acquisition processes in the classroom: Learning*

Japanese. Lawrence Erlbaum.

Rijlaarsdam, G., & Van den Bergh, H. (2006). Writing process theory: A functional dynamic approach. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *The handbook of writing research* (pp. 41 - 53). Guilford Publications.

Shortreed, I. M. (1993). Variation in foreigner talk input: The effects of task and proficiency. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 96-122). Multilingual Matters.

Soleimani, M., Modirkhamene, S., & Sadeghi, K. (2015). Peer-mediated vs. individual writing: Measuring fluency, complexity, and accuracy in writing. *Innovation in Language Learning and Teaching, 9*(3), 1-15. https://doi.org/10.1080/17501229.2015.1043915

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, *52*(1), 119-158. https://doi.org/10.1111/1467-9922.00179

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14*(3), 153-173. https://doi.org/10.1016/j.jslw.2005.05.002

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research* (2nd ed.). Sage.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Watanabe, Y. (2008). Peer-peer interaction between L2 learners of different proficiency levels: The interactions and reflections. *Canadian Modern Language Review*, *64*(4), 605-635. https://doi.org/10.1353/cml.0.0008

Wells, G. (2000). Dialogic inquiry in education: Building on the legacy of Vygotsky. In C. Lee & P. Smagorinsky (Eds.), *Vygostkian perspectives on literacy research: Constructing meaning through collaborative inquiry* (pp. 51-85). Cambridge University Press.

Yu, R. (2008). Interaction in EFL classes. *Asian Social Science, 4*(4), 48-50. https://doi.org/10.5539/ass.v4n4p48

Yule, G., & Macdonald, D. (1990). Resolving referential conflict in L2 interaction: The effect of proficiency and interactive role. *Language Learning, 40*(4), 539-556. https://doi.org/10.1111/j.1467-1770.1990.tb00605.x

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Iranian EFL Teachers' Conceptual Pedagogical Perception of the Role of Technology Integration in Classrooms

Research Article
pp. 111-136

Samira Arabzamani[1]
Ali Arab Mofrad*[2]
Ali Rajabi Esterabadi[3]

Received: 2021/02/01 Accepted: 2021/03/06

Abstract
Although extensive research has been conducted on successful technology integration, encompassing the readily available technology, providing a techno-based environment, and applying high-level technology in the classroom still seem to be very much in the minority. Moreover, technology integration illustrates additional barriers, including teachers' and students' technology-based knowledge and schools' financial problems. In contrast to earlier findings in English as a Foreign Language (EFL) contexts, however, little evidence seems to be available about Iranian teachers' classroom uses of the technology. Therefore, this paper aims to present the Iranian teacher's conceptual pedagogical perception of technology integration and illuminate how technology integration impacts teachers' classroom practice. To this end, 10 Iranian EFL teachers in private and public schools in Golestan Province were interviewed. The qualitative data of the interviews were recorded in English and transcribed; they were then coded and analyzed by employing thematic analysis (Braun & Clarke, 2006) via NVivo 12 computer software to show the dimensions and components of the conceptual model of technology integration. One of the themes extracted from the thematic analysis of the Iranian teachers' perceptions was the lack of technological knowledge. That is, a limited understanding of teachers about using technology in their teaching was viewed as a significant problem. Another theme referred to the limited use of technology which makes teachers use computers as a supplementary tool. In addition, it was found that the educational context does not create need-based technology conditions for them, leading to the teachers' ineffective use of technology.

keywords: teacher's perception, technology integration, thematic analysis, grounded theory, Iranian EFL teachers

---

* Corresponding Author
[1] MA Student, Department of English Language Teaching, Gorgan Branch, Islamic Azad University, Gorgan, Iran. teflsaz@gmail.com
[2] Assistant Professor, Department of English Language and Literature, Faculty of Humanities and Social Sciences, Golestan University, Gorgan, Iran. a.arabmofrad@gu.ac.ir
[3] BA Student, Department of English Language Teaching, Golestan University, Gorgan, Iran. alirajabi167@yahoo.com

DOI: 10.22051/LGHOR.2021.33557.1398

Introduction

The development of technology has garnered attention to the importance of integrating technology into English as a Second Language (ESL) or English as a Foreign Language (EFL) classes regardless of technical knowledge of teachers. However, the degree of employment of technology may vary according to teachers and students' respective perception and skill concerning the technological means. In this respect, some researchers asserted that the instructors' perceptions about technology integration are the vital elements illustrating the content and scope of the implementation of technology in the curriculum (Becker, 1991; Campoy, 1992; Ertmer et al., 1999; Pedersen & Liu, 2003). That is, as Marcinkiewicz (1993) and Miller and Olson (1994) noted, besides the teacher's belief, which is also a significant constituent for applying technology in lesson plans, the awareness of teachers is a vital factor. As a whole, teacher's perception is accounted as influential and valuable input for planning and development of technology-based materials and curriculum.

Many studies were conducted regarding the significant impact of Iranian EFL teacher's perception on the utilization of technology in classroom instruction (Ashrafzadeh & Sayadian, 2015; Gilakjani, 2017; Rahimi & Yadollahi, 2011). Some research studies in Iranian EFL context indicate that in contrast to student-centered teaching, traditional teacher-centered teaching does not achieve any successful utilizations of technology, (e.g. Amiri & Saberi, 2019; Lak et al., 2017; Zohrabi et al., 2012). Therefore, it seems essential to modify the conventional system of instruction to a modern and constructive one (Jonassen et al., 1999; Strommen & Lincoln, 1992). In addition, there are gaps related to technology integrations among Iranian EFL teachers (Rahimi & Yadollahi, 2011), one of which is inadequate digital content (Ertmer, 2005; Hew & Brush, 2007; Project Tomorrow, 2011). Moreover, as stated by Johnson et al., (2010), teachers' lack of access to the internet connection in schools creates another critical issue that negatively affects the progress of technology integration.

Therefore, this study aims to explore Iranian EFL teachers' perceptions toward the role of technology integration in instructional classrooms. Additionally, it is important to take the problems and issues that teachers confronted and also the effects of the integration of technology in the classroom context into consideration.

Review of the Related Literature

The Role of Technology Integration in Education

Correct training of teachers should encompass both the application of technology for personal needs and the use of technology with students in the classroom (Burkholder, 1995; Hignite & Echternacht, 1992; Hochman et al., 1993; Kearns, 1992; Ritchie & Wiburg, 1994; Todd, 1993; Wetzel, 1993; Woodrow, 1992). The generalizability of these results is that teachers ought to be computer literate and seek to include technology in their lesson plans; in contrast, teachers who do not have computer literacy and updated knowledge regarding technology do not achieve their goals. Such differences in the application of technology in the classroom play an important role in teachers' and students' perception towards technology integration.

Numerous findings and additional evidence contribute to the suggestion that there are crucial elements that directly and indirectly have a significant impact on technology integration. According to Howley et al., (2011), the term technology

integration has been influenced and facilitated by variables such as schools, teachers, and students. In this respect, Ritzhaupt et al., (2012) point out the existence of several practical factors. First comes teacher-related variables like teaching experience in an educational environment, and their level of education. Secondly, there are school variables that encompass the availability of technology and financial-technological professional support. Lastly, there are student-centered variables that include student's level of Proficiency and motivation, as well as the number of students in the class. In sum, the results suggest that teachers' willingness and readiness, as well as their experience with technology, have a considerable impact on the degree of integration of technology in their educational goals (Inan & Lowther, 2010; Ritzhaupt et al., 2012; Tondeur et al., 2008). In the light of Ritzhaupt et al. (2012), Inan and Lowther (2010) argue that two critical factors having the highest effect on technology integration are technology's accessibility and teacher's perception. In other words, teachers' perception toward using technology-based materials in the classroom was considered as the second most pivotal element (Inan, & Lowther, 2010; Tondeur et al., 2008).

The integration of technology has been widely investigated among Iranian EFL context during last decade (e.g. Mahmoudzadeh, 2014; Soleimani & Arabloo, 2018). According to Riasati et al. (2012), comprehending the benefits and barriers of technology integration is a significant step in the Iranian EFL educational system. In their study, it has attempted to present a variety of factors that encompass technology integration based on the perceptions of Iranian EFL teachers including engagement, collaborative learning, and improvement in academic ability which advocate the application of ICT in language teaching. As an example, Mahmoudzadeh (2014) studied on the effectiveness of Microsoft PowerPoint presentation and its role on the Iranian EFL learners' vocabulary enhancement. To this aim, 120 Iranian EFL participants were enrolled in a quasi-experimental design and the results revealed that using PowerPoint software creates an interaction among the learners which improved their scores in vocabulary tests. Additionally, Mahmoudzadeh (2014) stated that PowerPoint software offers several possibilities such as sounds, colors and actions which motivated students to integrate with vocabulary learning as much as possible.

Furthermore, the effectiveness of technology in the field of Iranian EFL teachers' training were investigated by Soleimani and Arabloo (2018). They investigated the influence of technology training courses on implementations technology in EFL classroom atmosphere by Iranian EFL teachers. The results of this study revealed that the teacher training courses held by the Minsitry of Education was not proper enough to improve Iranian EFL teachers in terms of real-life technology integration in EFL classroom context, as they noted that "was not at a proper level and was limited to the use of some common technologies such as printed materials, computers and LCD projects." (p. 14).

However, the lack of teachers' and students' knowledge, inaccessibility, and the lack of time imped the extensive use of technology in the Iranian EFL atmosphere. Accordingly, they noted that the awareness of positive and negative aspects of technology have a vital impact on practicing Iranian EFL teachers' perceptions. Indeed, awareness in this manner provides essential information about technology integration for interested Iranian EFL teachers.

Although having access to technology in schools has been facilitated during the last decade, reaching the goal of applying technology in classrooms is not

solely related on technology-related factors (Tondeur et al., 2016). In a longitudinal case study, Levin and Wadmany (2008) showed that teachers' views towards technology integration were shifted from supplementary materials to the needs of the classroom to the main course. Such modified views were constructed based on teacher's experiences in the educational environment. However, the popularity of technology does not show or prove the successful use of technology by teachers as an inseparable part of their educational instruction or in general. Teachers believe that the engagement and the appliance of technology into their teaching progress can be fruitful, though they mostly struggle with an integrated approach to balance the technology with their teaching curriculum (Dexter et al., 1999).

Computer-Mediated Communication (CMC) and Language Teaching

In an analysis of Computer Assisted Language Learning (CALL), Ahmad et al., (1985) found the term computer is used when referring to a useless device. However, according to Taylor's (1980) view, a computer is mentioned as a teaching machine in the classroom. Nowadays, computers are used to present materials, solve skill deficits, practice skills at their own pace, and foster skill fluency or automaticity (Roblyer, 2006). In this respect, Bodomo (2010, p. 6) defined CMC as "the coding and decoding of linguistic and other symbolic systems between sender and receiver for information processing in multiple formats through the medium of the computer and allied technologies such as PDAs, mobile phones, and blackberries; and through media like the internet, email, chat systems, text messaging, YouTube, Skype, and many more to be invented."

As it is evident CMC, encompassing social media, email, online discussions, Key pal, and internet-based applications, is a fundamental component of language learning that can fast become a vital instrument in language learning. The aim of CMC is to boost the level of students' communication and integration (Kern, 1995; Warschauer, 1997). Additionally, it decreases the sense of fear and shyness in discussions (Kelm, 1992; Kronnenberg, 1994). Moreover, in contrast to face-to-face communication, it gives plenty of time for reflection and thinking (Abrams, 2001; Duffy et al., 1996).

An overview of the related studies has revealed that the functions and interactions of computer programs have been expanded through the past two decades in EFL classrooms (Ahmadi, 2018); in this regard, Donaldson and Kötter (1999) and Lee (1998) noted that the employment of computers, internet-based applications, and software have changed the world of teaching. Endless uses of computers, for instance, monitoring learning, giving comments and feedbacks, and giving internet-based tests, have been developed beyond the teacher's expectations. These developments have considerably modified the functions of the computer as an instructional tool and changed it to a learning tool, consequently shifting the focus from the real world to the virtual world. The emergence of computers, software, and online-based applications, such as learning tools, has helped exposing students to authentic data. In the same vein, Warschauer and Healey (1998) noted that multimedia and advanced techniques of the utilization of computers have provided authentic information and have also exposed students virtually, thus leading to emersion of integrated language skills in students. In addition, a variety of internet-based software have been introduced as supplementary material to teachers that boosts autonomous learning. Moreover, CMC, despite the practical constraints, has made a significant contribution to task-based activities, further boosting peer work

and group work activities (Crook, 1994; Kern, 1996; Warschauer, 1995) and enhancing motivation and collaborative learning (Braine, 2004; Hanson-Smith, 2000; Ushioda, 2000).

Generally speaking, many studies in the field of CMC have only focused on the motivational aspect of the use of computer. That is, the computer is considered as a motivational tool for cooperative and communicating aims. According to Vygotsky (1986), the use of CMC and CALL outweighs communications and increases student's motivation. To be crystal clear, it could be mentioned that there is a vital requirement for noticing psychological benefits in CMC that foster with factors such as motivational aspect during cooperative tasks via computers. The application of technology for language instruction highlights an improvement in communicative skills among Iranian students with the mediated role of computers that benefits from motivational aspect (Karimi et al., 2014). Therefore, CMC or CALL promotes and assists the cooperative sense of an individual; in other words, each individual can construct and expand their knowledge with others which can be called a socially-mediated process.

However, a number of significant limitations need to be considered. First, the lack of feedback (Salaberry, 2000), second, the shortage of nonverbal cues (Ferrara et al., 1991), and third, the lack of learner's first language (Lee, 2004) tend to highlight the inattention to the use of computers in the educational classroom. Another problem regarding the use of computers is that it fails to take pedagogical approaches, language skills, learning styles, students' target language proficiency levels, and motivation into account. As Warschauer (1996) stated, computers can enhance authentic communication and promote student-centered teaching practice.

Teacher Perceptions toward Technology

According to Fishbein and Ajzen (1974), the term ‘attitude’ generally refers to the predispositions of a person to favorable or unfavorable feelings toward people, events, and objects. Teachers' and students' attitudes seem to play a prominent role in reaching educational goals and may also act as an essential factor in an educational environment leading to the integration of innovation in an educational setting (Albion & Ertmer, 2002; Becker et al., 1999; Pajares, 1992). Overall, there are many challenges when it comes to its definition, as the Fishbein and Ajzen’s (1974) is regarded as the major approved definition in this study. A broader perspective has been adopted by Pajares (1992), who labelled the concept of a teacher's beliefs as a "messy construct" (Pajares, 1992, p. 307).

It is essential to know to what extent teachers integrate technology into their teaching practices. Studies have shown that using technology in a learning context can be a complicated process (Scherer et al., 2018). The use of technology among teachers has dramatically increased; however, their perceptions toward technology are continuing to be a problem for educational institutions (Bishop & Spector, 2014). In this respect, while most studies in the field of teachers’ perceptions have only focused on negative views concerning technology, it is considered as a powerful instructional technique (Clark, 2000). Recently investigators have examined EFL teachers' attitudes toward instructional setting (Kollias et al., 2005). Factors found to be influencing a collaborative learning system encompassed both changing in technology as well as shifting teacher's perceptions. In fact, positive teacher attitudes are quickly becoming a vital instrument in the integration of computer use and also in shifting teachers' negative

views on technology (Kollias et al., 2005).

Cox et al.'s (2003) study of technology provides an exciting opportunity to advance and improve our knowledge of the application of technology in the instructional environment in order to modify teaching strategies. They found out that teacher enthusiasm can have vital effects on technology adaptation in the classroom. This was also consistent with the results obtained by Clariana et al. (1991) and Sepehr and Harris (1995). Clariana et al. (1991) found out that the missing link in educational system's software planning for teachers were mostly ignoring teachers' attitudes. Also, Sepehr and Harris (1995) noted that the preferred choices of teachers could improve the use of drills and practices of class books among them. In addition, the impressive surveys of Coniam (2002) and Ducate and Arnold (2006) showed that instead of applying technology solely to teach in an instructional environment, teachers could also learn with technology. Similarly, Hadley and Sheingold (1993) indicated that teachers should make a commitment to foster their professional development and leaning knowledge to integrate technology and innovative tools in educational settings.

Habibi et al. (2019) investigated the effectiveness of EFL teachers' perceptions toward implementing technology in their classrooms. To this aim, 138 pre-service EFL teachers were enrolled in this study to answer a researcher-made survey and the results indicated that most of Indonesian EFL teachers who participated in this study had positive perception toward the use of medium and advanced ICT in their EFL classrooms. On the other hand, it was found out most of the participants believed that the ICT integration in their classrooms faced with lack of implementations in feedbacks and peers-collaborations.

Therefore, Marcinkiewicz (1993) pointed out that adopting technology in the classroom is fostered through modifying teacher's perceptions. Similarly, this view is supported by Violato et al. (1989). Hignite and Echternacht (1992) asserted that both teachers' perception and teachers' experiences play an essential role in promoting the incorporation of computers in the classroom. So far, two factors have been identified as being potentially important; positive perception and computer literacy of teachers, the existence of which would lead to technology integration. Given all that has been mentioned so far, Woodrow (1990) suggested that the teacher's perception has a drastic effect on teaching children regarding technology integration. Todman and Dick (1993) also highlighted that teacher's perception towards technology integration leads to fostering the quality of the children's experiences of the computer.

Several studies thus far have linked student's perceptions with student's attainment, some of which concluded that there is a significant positive correlation between student achievement and student's positive perceptions (Office of Technology Assessment, 1995).

Research Questions

1. What are the Iranian EFL teachers' perceptions toward technology integrations?

2. What are the barriers to implementing technology integration among Iranian EFL teachers?

3. How are Iranian EFL teachers affected by technology integrations in EFL classrooms?

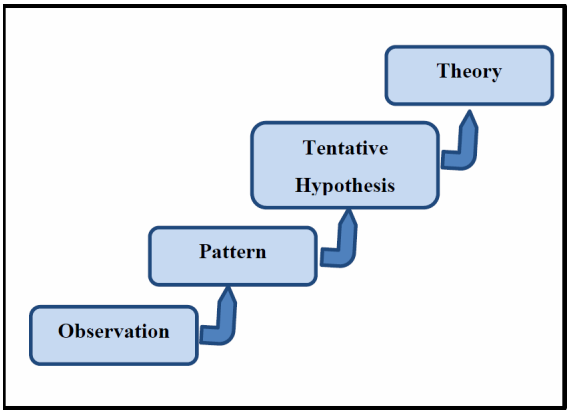
Methodology

Research Design in Grounded Theory

Grounded theory is a common design to perform in qualitative research design in order to generate theory and form data. The grounded theory is notionally flexible but can be a complicated methodology at the same time. In the present study, grounded theory, moving from data to theory, was applied as a foundation of research design. This methodology uses inductive method for qualitative study, and it begins with a question, as researchers review the data and ideas collected and tagged with *codes*, which have been extracted from the data. The collected and coded data can be grouped into concepts, and then into categories. These categories may become the basis for a new theory. Glaser and Strauss (1967) indicated that the inductive approach is the most suitable way for emerging theory from social data. This is highlighted in the grounded theory model, illustrating that whether the theory has a relationship with an empirical situation or not. As indicated in figure 1, and since the nature of the study is exploratory, an inductive approach is adopted in this present study to analyze the data; the researcher firstly gathered the information from the interviewees' observations based on which the theories were chosen.

Figure 1

Schematic Representation of Inductive Reasoning, Taken from Trochim & Donnelly (2006, p. 17)



Participants

The participants of this qualitative study, as indicated in Table 1, were 10 English language teachers (five males and five females), ranged in age from 26 to 47, who voluntarily took part in this study. They majored in English language teaching at BA (N = 1), MA (N = 6), and PhD (N = 3). Three of them were university professors, two of them were teachers of public high schools, and five of them were teachers of private high schools and language institutions. The participants' mother tongue was Persian and their teaching experience varied from 5 to 30 years. They were selected based on purposive convenience sampling owing to the lack of the availability of enough participants. Their background knowledge of computers was also different; however, most of them had ICDL 1 and 2 certificates. According to Amado-Salvatierra et al. (2016), International Computer Driving License (www.ecdl.com) is a series of certificates which mainly focus on basic computer and digital literacy such as online and word processing fundamentals.

Table 1
The Demographic Information of Interviewees

Name	Age	Gender	Qualification	Teaching Experience	Teaching Context
Teacher 1	43	Male	M.A	30	University
Teacher 2	47	Male	PhD	25	University
Teacher 3	39	Female	PhD	22	University
Teacher 4	32	Female	M.A	15	High School
Teacher 5	32	Female	M.A	12	Private Institution
Teacher 6	29	Male	M.A	11	High School
Teacher 7	30	Female	M.A	10	High School
Teacher 8	29	Male	PhD	10	Private Institution
Teacher 9	26	Male	M.A	7	High School
Teacher 10	26	Male	B.A	5	Private Institution

Instruments

In order to comprehend EFL teachers' perceptions of the application of technology, an open-ended interview developed by Kim (2008) was adapted and applied. The original version of the interview consisted of 27 questions to which the researcher added another five questions (questions 28 – 32) in order to include the barriers with which Iranian EFL teachers confront. Subsequently, the content validity was checked and reviewed by two scholars in this field. It includes items about general information, the experience of teaching, instruction and motivation strategies, teachers' perception, and teachers' barriers in implementing technology in the classroom. To illuminate the conceptual map of ideas, the statements of 10 participants were recorded and then transcribed in detail for further analyses. The analysis of the open-ended interview data was carried out through the use of Nvivo 12 to extract the codes. In addition, the exploratory approach of thematic analysis was employed to analyze and decode the data, to highlight the themes and to illuminate relationships (See Figure 2). It should be noted that, the two scholars were full time professor in Golestan University, Gorgan, Iran. They have studied applied linguistics and TEFL for over 10 years and written several scholarly articles and books. Accordingly, in order to check the inter-rater reliability, the Cohen Kappa was applied and the results revealed that the two scholars had agreement on their judgement on the interview as an instrument. There was fair agreement between the two scholars, $K = .733$ (95% CI, .300 to .886), $p < .0005$.

Procedures

First, the interview questions by Kim (2008) were given to the participants in advance, and they were given a chance to ask any questions before taking part in the study. In the interview session, while participants conveyed their perceptions and thought in regards to the role of technology in classrooms in English, their statements were recorded by a digital recorder. Three of them were interviewed on the phone due their inability to have face to face connection. Each participant was interviewed for 30 minutes. The recorded interviews were transcribed by transcription websites (e.g. http://www.amberscript.com). It should be noted that the participants voluntarily enrolled in this study and signed a constant form; additionally, they were ensured about the anonymity and confidentiality of the data.

Results

This study promised to examine teachers' beliefs about computer use in the educational environment. To present and analyze the data, the Grounded Theory was

adopted in the present study. Employing thematic analysis, as indicated in Table 2, the data were coded into six groups, namely, teacher's perception of computers, kinds of skill, kinds of materials, kinds of facilities, the experience of using computers, and barriers of using computers.

Table 2

Thematic Analyses' Descriptions of Participants' Interviews

Theme	Node/Theme
Interviewee 1	barriers of using computer and technology, lack of facilities in the university, school, institute, lack of knowledge of students, lack of knowledge of teachers, Experiences of using computers, having the certificate of computer, lack of a certificate of computers, Kinds of facilities Internet-based application, Social Media
Interviewee 2	Experiences of using computers, having the certificate of computer, Kinds of facilities Internet-based application, Social Media
Interviewee 3	barriers of using computer and technology, lack of facilities in the university, school, institute, Social Media, Experiences of using computers
Interviewee 4	barriers of using computer and technology, lack of facilities in the university, school, institute, lack of knowledge of students, lack of knowledge of teachers, Experiences of using computers, having the certificate of computer, lack of a certificate of computers, Kinds of facilities Internet-based application, Social Media
Interviewee 5	barriers of using computer and technology, computer, Kinds of facilities Internet-based application, Social Media
Interviewee 6	barriers of using computer and technology, lack of facilities in the university, school, institute, lack of knowledge of teachers, Experiences of using computers, having a certificate of computer, lack of a certificate of computers, Social Media
Interviewee 7	technology, lack of facilities in the university, school, institute, lack of knowledge of students, lack of knowledge of teachers, Experiences of using computers, having a certificate of computer, lack of a certificate of computers, Kinds of facilities Internet-based application, Social Media
Interviewee 8	barriers of using computer and technology, lack of facilities in the university, school, institute, Experiences of using computers. Lack of the certificate of computers, Kinds of facilities Internet-based application, Social Media
Interviewee 9	barriers of using computer and technology, lack of facilities in the university, school, institute, lack of knowledge of students, lack of knowledge of teachers, Experiences of using computers, having the certificate of computer, lack of a certificate of computers, Kinds of facilities Social Media
Interviewee 10	barriers of using computer and technology, Experiences of using computers, having a certificate of computer, lack of a certificate of computers, Kinds of facilities Internet-based application, Social Media

Based on open and axial coding of the qualitative data of interviews using NVivo12, the major theme of *Teacher's perception of technology* emerged which included some sub-themes with either a direct or indirect relationship to a major theme. According to the results of open and axial coding of interviews, as indicated in Table 3, six categories has been identified as the main categories of technology integration. The sub-themes were categorized into *kinds of materials, kinds of facilities, skills, experiences of using computers, and barriers of using computers and technology*. The contextual conditions are factors like the lack of facilities in universities, schools, and institutes, which have a direct relationship to the major theme. In addition, causal conditions that have a direct effect on the major theme ''teacher's perception of technology" encompass the lack of teacher's knowledge, skills as well as the experience of using technology. Therefore, they are the first and most important factors that have a significant impact on teacher's perceptions. In general, kinds of skill, facilities, materials are called intervening conditions out of the teacher's control. The rules have an indirect effect on "teacher's perception of technology". Therefore, these rules are not without any effect. The "skill" theme is called the *mediating condition* since it mediates between two themes and has a direct effect on the major theme (i.e., teacher's perceptions of using technology) and a sub-theme (i.e., kind of facilities).

Table 3

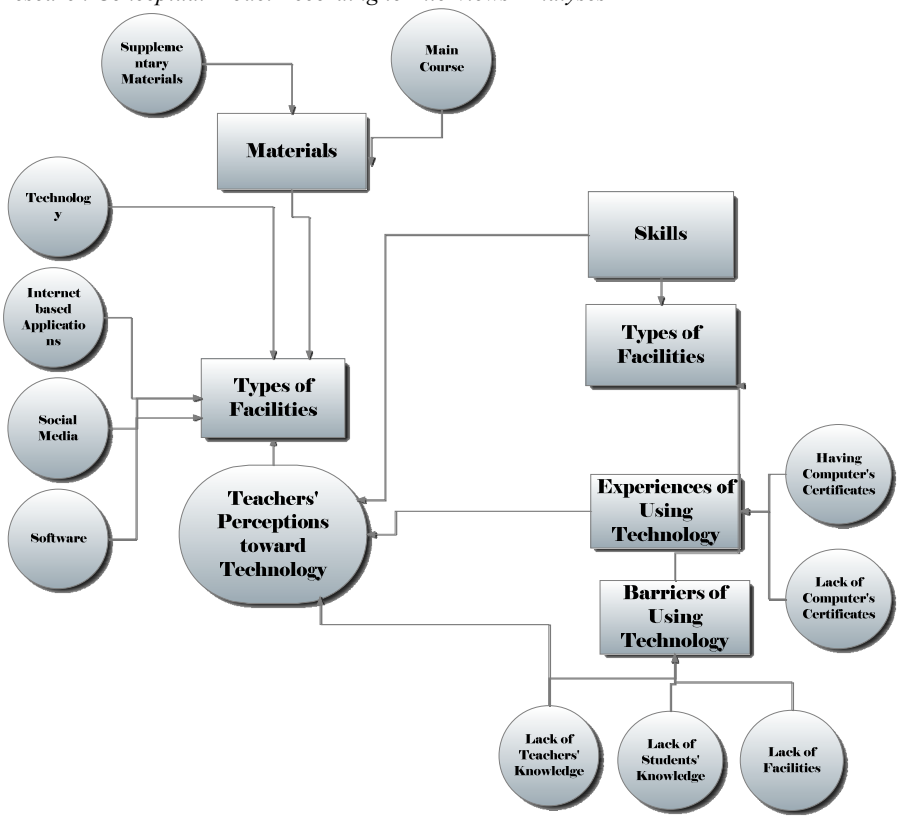
Categories and Subcategories of Frequency of Technology Integration

Description	Interviewers	Frequencies
Barriers to Using Computer and Technology	10	29
Lack of Facilities in Universities, School, Institute	6	15
Lack of Knowledge of Students	2	2
Lack of Knowledge of Teachers	6	13
Experiences of Using Computers	7	7
Having a Certificate of Computer	5	5
Lack of a Certificate of Computers	2	2
Computer-Mediated-Communication	0	0
Internet-Based Application	9	61
Social Media	8	37
Software	8	35
Technology	10	378
Kinds of Materials	10	22
Main Course	3	3
Supplementary Tool	9	19
Skills	10	49
Teacher's Perception of Computer	10	87
Positive Perception	7	75
Negative Perception	3	11

Figure 2 depicts a more vivid thematic map of the data and conceptual understanding of all nodes.

Figure 2

Research Conceptual Model According to Interviews' Analyses



Responses to Research Question One

1. What are the Iranian EFL teachers' perceptions toward technology integrations?

The participants' views showed different themes for the concept of technology integration. In this respect, a variety of comments on technology integration can be described under positive and negative views.

Negative Perception. In line with the integration of technology in the educational environment, the teacher's perception is considered as the essential fulcrum that has a pivotal effect on teaching. CALL in Iranian EFL classrooms illustrated several negative thoughts among Iranian EFL teachers. We would first like to highlight a negative theme raised by two participants (Interviewee 10 and 3); the conversation was as follows: "a good teacher is a teacher who can teach without any technology or computer." This teacher believes that teaching language is an interactive activity, one that is facilitated through face to face conversations since a conversation includes lots of hesitation, false start, and pauses as well as feedbacks given to interlocutors. Then, the time of conversation is short; it means, speakers do not have time to think. A speaker has to control thinking and emotion at the same time. In other words, using technology in an instructional setting cannot create a real environment at all. Other interviewee mentioned that "I believe that technology is only a machinery, a distracting tool. This is while language is a kind that, of course, must be communicated first via face-to-face means". They believe that although

well-established technologies are maturing, such as mobile devices, internet access, and smart boards that are accessible and available for foreign language learning, these innovations would just lead to an increase in learner's interest and motivation; however, the innovations provide inappropriate, shallow interactional opportunities and inaccurate comments leading to distracting learners from the learning environment.

Positive Perception. It can be indicated that the use of technology has a negative and positive impact on education. In other words, access to technology is not the only criterion. Technology integration should have been applied in a way that facilitates learning. In this regard, eight out of ten teachers interviewed would like to integrate computers or technology into their materials; however, it appears the use of computers is bound up with teacher's perception, computer's skill, and knowledge. Therefore, most of the teachers said the cooperation of school and universities is an influential factor in the application of technology in teaching curriculum. This theme was presented by one of the participants: "I gain many benefits from using the computer to facilitate my process of learning." Based on this teacher' comments, the application of technology in instructional setting make student's motivation increased, and most of the silent student participate in learning, and therefore, they attain confidence.

In this regard, interviewee 4 indicated :"I think the use of computers and the internet in the classroom is one of the fruitful things that has happened in the last century of teaching and I use the computer and the internet in my classroom most of the time as the main instruments. The creative role of the teacher is to update her/his method and create a new strategy for teaching. As times goes by, students become more interested in technology-based life; therefore, it is necessary to create computer-mediated communication. I employ Pen Pal-related activities for discussions and sometimes interact with students through WhatsApp. Students show enthusiasm to use social media in learning a language". In addition, the evolution of technology is beneficial to students for several reasons. It increases the efficiency of the teaching and learning process, and students can communicate with anyone who happens to be geographically speaking far. It also encourages individual learning and collaboration.

Responses to Research Question Two

2. What are the barriers to implementing technology integration among Iranian EFL teachers?

Lack of Teacher's Knowledge. Issues like low internet access, lack of alignment between teacher's instruction and technology, and inconsistent computer availability of teachers make it difficult to use technology for teachers who are not capable of updating their knowledge. The main reason for teachers' rejection of integrating technology in the classroom comes from the lack of awareness and the lack of the needed knowledge to implement technology with their instruction. Therefore, there is little motivation for teachers to adopt technology in classrooms. As interviewee 5 indicated, the lack of technical and theoretical knowledge is another barrier to the use of Computer-assisted Language Learning technology. Not only is there a shortage of knowledge about developing software to promote learning, but many instructors do not also understand how to use the new technologies. According to the conceptual map, the lack of teacher's knowledge is considered as a fundamental factor that has a direct effect on the main theme. For

instance, interview 2 said "as you know, since schools, universities, language labs do not have the needed facilities, teachers' knowledge cannot be updated regarding technology-related instruction and, most of the time they left behind. They are not very advanced in this case".

Additionally, as indicated in Table 3, six interviewees talked about the lack of teacher's knowledge as the most apparent barrier. For instance, Interviewee 9 believed that "teachers go to computer-centered classes but they do not learn about how technology should be employed in their teaching process, how to use computers, how to use PowerPoint or how to connect to the internet". This problem was especially highlighted by the experienced teachers who had more than 10 years of teaching experience more than novice teachers. Such old and experienced teachers seem to be more in need of updated knowledge of teaching. Although technology cannot replace teachers, it provides new opportunities for enhancing student's knowledge. Technology-assisted learning seeks to remedy the problems of the educational process as well as facilitates the student's process of learning.

Lack of Student's Knowledge. As interviewee 4 put forward, "Encouraging creativity in the classroom is the main task of a teacher which powers their divergent thoughts with the ability to use computers and online tools. Providing opportunities for students to improve their language process through computers should be the key solution of classrooms' issues. Some teachers may present these chances through tasks or group activities where every student's contribution is valuable to the development of their creative thinking as students attempt to describe a solution to a problem to another member of the group. As many students do not have facilities at home or do not entirely know how to use technology, they do not cooperate with teachers at all. In addition, families disagree with the application of technology, due to their belief that technology distracts students when learning. "Their technology knowledge is obsolete". He means that human beings live in a time that changes occur fast. The changes move rapidly and inhibit the appliance of conventional norms; as a whole, it will provide new opportunities for the future. Despite these rapid changes, integrating technology-assisted learning is an important issue that needs time and willingness of parents and students. As Interview 8 put it: "Not all students have any cellphones or laptops at home, moreover some families completely disagree with the students' use of technology or social media. In addition, most of the students' computer knowledge is low". As table 3 reveals, these two interviewees believed that the main factor that makes students be behind is their parents. Moreover, as a conceptual map illustrates, the lack of students' knowledge is subcategories of barriers of application of technology in the instructional setting.

Lack of Facilities in the Institute, School, and University. Barriers of technology integration have been categorized differently. Lack of facilities in instructional environment is considered as a vital factor. As interviewee 2 stated "I told you schools, universities, language labs do not have the particular facilities, do not have technical facilities. They cannot provide them for teachers and students because of the financial problem; However, I have to keep up with a technology-related instruction". He also mentioned that further research should be done to investigate the financial barriers. The cost of hardware, software, and staff development are high. Therefore, it can be stated that the cost could be a significant factor in technology development. Financial problems make university or school do not collaborate with teachers, and they cannot pay attention to the need of students and teachers to technology. As ideas extracted from interviewee 1 shows, "As I

heard from my colleagues, schools and institutions come up with the financial issue, so they cannot provide a technology-based setting for their students". She also added that successive increases in the intensity of the budgeting policies of universities is one of the significant issues that instructors encountered. In other words, several factors that inhibit use of technology in the classroom include low computer experiences, teachers' the lack of knowledge, the lack of computer access, and the lack of access to resources. So these issues might lead to teachers, demotivation in implementing technology in classroom.

As interviewee 9 believed, "There some students have now limited access to the internet in the region they are living. Most schools do not provide any technical equipment for teachers. They prefer budgeting on important resources at school instead of technology". He explained that it should be noted that now, the cost of technology-assisted learning might be conceived as high. Therefore, if the budgeting of technology-assisted learning is considered, it leads to the higher success of educational setting and instructional time. As Table 3 indicates, lack of facilities in the educational setting has the highest frequency in comparison with other barriers. It means that it is so hard without technical support teachers can overcome barriers of integrating technology. These technical failures make teachers resistant to new changes

Responses to Research Question 3

How are Iranian EFL teachers affected by technology integrations in EFL classrooms?

The Technology was Used as the Main Course. Regarding the role of computer technology, all teachers stated that technology integration is beneficial for language learning; however, teachers had varied background experiences, so their aims and perception varied as well. In this respect, a variety of perspectives was expressed in these ten teacher's perceptions of technology, and there were two emerging themes: employing technology as the main course or as the supplementary course. When the teachers were asked why they seek to apply technology, five teachers stated since technology expose the students to authentic text and meaningful communication, it acts as a motivator so that it ought to be taught as a mandatory course in the classroom. Therefore, the teacher must update their professional technological skills. In other words, technology integration foster student-centered learning. It means, in an instructional classroom, students themselves choose their learning materials, can make decisions on what kinds of skills must be improved and can adjust the speed of learning based on their own needs. Generally, it fosters meaningful learning. As Interview 4 noted, "The use of technology in the class makes students especially kids and teenagers motivated to have more eagerness to learn a language.

Moreover, it helps them to be a better listener by providing them with authentic inputs, especially making them be familiar with new words and be familiar with the British and American accents. It also improves their speaking so they can find themselves motivated to speak and find new vocabularies and idioms that they need". Also, Interview 10 mentioned that "Let me have two types of teachers. Teachers who teach in state or private schools. You should especially teach teachers in public schools to use a computer the Smart board and projector in class and They must use them as a main course. To me, the use of the internet is not just as a teaching aid to solve the problems, but it must be seen as the main course".

The Technology was Used as a Supplementary Course. In contrast to teachers' voices above, one of the teachers argued that technology integration ought to be considered as having the peripheral role. More specifically, he believed that one intriguing feature of the teacher is teaching English without including technology in lesson plans. Moreover, four teachers believed that technology integration is an optional or supplementary tool, and technology's role is limited to particular contexts and language skills. That is, they emphasized teacher-centered learning. For instance, Interviewee 6 stated that "Yes, I use the computer only to play media or educational movies or to show students some pictures related to their lessons. I use technology as a supplementary instruction". As interviewee 8 elaborated: "Policymakers and scholarly society are not ready for such developments and changes in teaching. In Iran, using social media or computers depends on the knowledge and motivation of teachers. In fact, due to the lack of teacher's knowledge, they use technology and computers as a supplementary course.

Technology is a tool that can modify the system of teaching and learning. Today instructors want to implement technological tools in their teaching in order to facilitate level of learning. In a technology-rich classroom, technology just provides authentic conditions for learning. It does not mean students must learn technology. It is a means, not an end".

How Technology was Integrated into the Classroom. The point of view of teachers concerning technology in the educational setting, some factors that constrained technology integration in the classroom are content, skills, time and frequency. Interviewee 7 elaborated her restriction in applying of technology on the content and skill: “Technology-related softwares need special techniques. Some of them are very technical so that I use some parts of them related to my lessons. I myself cannot use them for all skills or lessons. I am not very technical. Teachers have to somehow make students learn to employ technology”.

In fact, as the online databases of English, for translating Persian to English or English to Persian while I am teaching, for example, I personally use Parse online translator, Google translate, and Trados. This can be really fruitful for getting the meaning across at the time of story-telling or reading comprehension in my classes. Therefore, for translating a text, I can use the computer as an assistant. In addition, we can use a computer for writing a summary of that text on Microsoft Office Word”.

Interviewee 2’s comment shows how to plan the use of technology for interpreting texts. Although she is not a technical teacher in order to integrate technology for all skills evenly, she believed technology fosters writing skills. Overall, Teachers’ perception on the application of technology contradicts entirely the point of views of integrative application of CALL (Warschauer & Healey 1998) and constructive strategies of applying of technology (Jonassen et al., 1999). As a whole, as mentioned, the implementation of technology was restricted in term of time, frequency and skill.

Discussion

The findings of this study confirm and extend the indications of previous literature about the complications of integrating technology into classrooms despite the positive aspects that most of the candidates mentioned. As stated by Bishop and Spector (2014), teachers’ perception is a key factor to consider in regards to the integration of technology. While the majority of the interviewees believed in the

efficiency and the convenience of technology, there was a general consensus that functionality can only be achieved with adequate knowledge, perception, and skill. Technology integration seems to highly influence the teacher's perception about students' learning. They believed that technology-based classrooms motivate and foster students' computer literacy and enjoyment. Also, other positive outcomes of technology integration are the increase of student-centered learning and the development of four skills simultaneously. However, there are some concerns about barriers of applying technology in the classroom including applying low-level technological means like composing text in Microsoft Word, checking the vocabularies' translations in dictionaries on cellphones, playing music and movies, and presenting lessons by Microsoft PowerPoint.

Moreover, according to our findings, some believed that the use of technology is in contrast with the traditions of teachings, making the whole process less authentic as, according to their beliefs, teaching must be done via face-to-face means. They also claimed that the integration of such means would dismiss the importance of contextualization due to the limiting factors when it comes to emotions and intellectual processes. Teachers' knowledge also played a crucial role based on our findings which explains how the lack of the theoretical necessities would result in the rejection of technology. This lack of knowledge also concerns students as they will also be asked to incorporate such means.

Additionally, the application of technology and the internet in the classroom were controlled and confronted with some barriers like the lack of budget, facilities, and educational setting rules being out of teachers' control. Also, some teachers mentioned that ethical considerations of parents and students lead to the diminishing of the teacher's interest. While many of the families believed in the benefits of the integration of technology, some view it as a detrimental tool that would distract students when learning due to the many layers the utilization of technology has other than the educational ones.

The findings of this study show that technology-integrated education plays a significant role in instruction for Iranian EFL learners. Another crucial issue is how to apply technology to education. Technology alone does not make a difference for attaining such desirable outcomes, and teachers should also consider the variety of technology-based pedagogies as a benefit for related content, too, as it is noted in several studies like Jimoyiannis (2010), Koehler and Mishra (2009), and Usluel et al., (2007).

Jimoyiannis (2010) studied the perceptions of science teachers' program regarding the integration of the ICT in a classroom context. In Greece, there was a program called Technological Pedagogical Content Knowledge (TPACK) which was questioned from science teachers. They interviewed an overview on science instructors' program on teachers' perceptions. They found out that there were successes but they also met some challenges on the way of using technology integrated programs in the classroom which shows positive understandings with this study and the current study. Also, Koehler and Mishra (2009) had a similar study on TPACK. They interviewed the teachers' perceptions toward technology. They found out that the challenges divided into three frameworks into content, pedagogy and technology. Therefore, it can be stated that the findings of this study is in the same line with the current study.

Furthermore, Raygan and Moradkhani (2020) had a related study on TPACK. They examined the influence of TPACK, school climate, and perception on

the success in technology integration in EFL classes with the use of data collected from 209 Iranian EFL teachers. They found out that perception had a significant impact on technology integration, concluding that positive perception would contribute to an effective integration. To attain this perception, teachers must become familiarized with technology and become experienced. In addition, the accessibility of equipment is also a major factor to consider in regards to such perception. Nevertheless, TPACK, as stated by the authors of this study, remains the most powerful predictor for technology integration, meaning that without sufficient TPACK, teachers may still feel uncomfortable with the concept of technology integration, which further supports the findings of the present study.

Another related study was investigated by Usluel et al. (2007). In this study they studied on ICT with regards to different affective factors such as level of education and age on 590 teachers in Ankara. They gathered the information through a developed questionnaire to understand the amount of different technology integrated applications used such as word processors, search engines, databases and the Internet Explorer. They grasped that there were some obstacles such as lack of internet connection and the absence of teachers' knowledge concerning of how ICT can be implemented in language teaching context. Finally, in order to overcome this obstacle, teachers must be trained and supported. Providing the conditions of needed-based technology for them is constructive. So these conditions make teachers take account of the role of technology integration in their own desired outcomes.

Last but not the least, Ding et al. (2019) studied on Chinese EFL teachers' perception toward technology integrated classes on 12 secondary-level EFL teachers on three dimensions namely skill-based, rule-based and function-based. The findings revealed that most of teachers dealt with issues such as technological content-specific materials. Also, this study supports the findings of the current study.

Conclusion

To conclude, the barriers toward using technology in the classroom can be divided into different stages, such as having the proper knowledge to use technology among EFL teachers, and it should be noted that this issue rooted in educational process in universities. On the other hand, the lack of technological equipment can be another barrier for EFL teachers. Therefore, they cannot shift from the low level of technologies to high levels owing to their lack of enough technology-based knowledge, it means they cannot overcome the technical obstacles of technology integrations.

The findings of this study could theorized that (1) pedagogical beliefs (2) technological knowledge and (3) technological facilities were thematically presented with respect to the participants' interviews. That is to say, the following illustrations have shown that integration of technology in Iranian EFL context is related to these three variables.

Data analysis revealed that the personal attributes and skills could be vital stages of technology integration in Iranian EFL context. Similarly, the lack of appropriate and useful training and professional development opportunities for Iranian EFL teachers and, the lack of outside support, teachers were concerned about selecting and integrating technology with the administrators' pedagogical beliefs. Accordingly, these lack of training leads to utilization of non-practical and knowledge-based content. This provides several barriers on application of reliable

and useful technology integration. Also, based on the participants' perceptions, along with the lack of technological facilities, these three variables were mentioned as the major obstacles. Finally, based on the following statements, it could be stated that pedagogical beliefs, technological knowledge and facilities creates an advanced technological integration in EFL classrooms.

As for the implications, the following study had some implications for educational decision makers, textbook authors and EFL teachers. Accordingly, based on the results of this study, if decision makers in Ministry of education paid more attention on understanding a more interactive use of technology for teachers, there would be less issues for teachers in terms of applying technology in their classroom. Also, educational textbook authors could use the results of this study to find out what an EFL teacher thought of the interference of traditional teaching (textbooks) and the modern version of teaching (using technology). Finally, Iranian EFL teachers could benefit from the results of this study since this study revealed the in-depth perceptions of Iranian EFL teachers in terms of using technology and their solutions to deal with the in-class technological problems which could pave the way for them to enhance their classroom's experience of applying technology.

This study also encountered some limitation during its administration. Among the main limitations of this work, the use of interview which was adapted by the researchers to collect information could be noted as limitation of the study since the use of interviews has some shortcomings in quantitative measurement. The second main limitation is the representativeness of the study samples. The teachers who participated in this study did it were chosen based on the convenience sampling, so there was no random selection involved to ensure that the sample was representative of the study population. Therefore, caution is needed when generalizing these results. The delimitations of the study can be divided into two main categories. First, in order to make it feasible, Iranian EFL teachers who had basic understandings of technology integrated in language teaching were chosen to participate in this study. Second, although the usage of technology in language teaching can be examined in various variables of teaching context, the current research only studied teachers' perceptions toward technology.

Suggestions for Further Research

This study followed the qualitative designs to find out the answer to the research question, another study with quantitative design can help the scholars and teachers to find out numerical understandings on the perceptions of Iranian EFL teachers toward technology integration in Iranian EFL teaching context.

References

Abrams, Z. (2001). Computer-mediated communication and group journals: Expanding the repertoire of participant roles. *System, 29*(4), 489-503. https://doi.org/10.1016/S0346-251X(01)00041-0

Ahmad, K., Corbett, G., Rodgers, M., & Sussex, R., (1985). *Computers, language learning, and language teaching*. Cambridge University Press.

Ahmadi, M. R. (2018). The use of technology in English language learning: A literature review. *Journal of Research in English Education, 3*(2), 115-125. http://dx.doi.org/10.29252/ijree.3.2.115

Albion, P., & Ertmer, P. A. (2002). Beyond foundations: The role of vision and belief in teachers' preparation for the integration of technology. *TechTrends, 46*(5), 34-38. http://dx.doi.org/10.1007/bf02818306

Amado-Salvatierra, H. R., Hilera, J. R., Tortosa, S. O., Rizzardini, R. H., & Piedra, N. (2016). Towards a semantic definition of a framework to implement accessible e-learning projects. *Journal of Universal Computer Science, 22*(7), 921-942. http://dx.doi.org/10.3217/jucs-022-07-0921

Amiri, F., & Saberi, L. (2019). The impact of the learner-centered approach on learners' motivation in Iranian EFL students. *International Academic Journal of Social Sciences*, *6*(1), 155-165. https://doi.org/10.9756/IAJSS/V6I1/1910015

Ashrafzadeh, A., & Sayadian, S. (2015). University instructors' concerns and perceptions of technology integration. *Computers in Human Behavior, 49,* 62-73. http://doi.org/10.1016/j.chb.2015.01.071

Becker, H. (1991). How computers are used in United States school: Basic data from the 1989 I. E. A. computers in education survey. *Journal of Educational Computing Research, 7*(4), 385-406. https://doi.org/10.2190/P2UT-R3U3-FK1L-B89L

Becker, H., Ravitz, J., & Wong, Y. (1998). *Teacher and teacher-directed student use of computers and software: Teaching, learning, and computing* (Report No. 3). National Science Foundation, Arlington, VA.; Office of Educational Research and Improvement (ED), Washington, DC. https://eric.ed.gov/?id=ED437927

Bishop, M. J., & Spector, J. (2014). Technology integration. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop, *Handbook of research on educational communications and technology*, (pp. 817-818). Springer.

Bodomo, A. B. (2010). *Computer-mediated communication for linguistics and literacy: Technology and natural language education*. Information Science Reference.

Braine, G. (2004). Teaching second and foreign language writing on LANs. In S. Fotos & C. Browne (Eds.), *New perspectives on CALL for second language classrooms* (pp. 93-107). Lawrence Erlbaum Associates.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. http://dx.doi.org/10.1191/1478088706qp063oa

Burkholder, J. (1995). *An annotated bibliography of the literature dealing with teacher training in the uses of the computer in education* (Report No. 3). Office of Education, US Department of Health, Education, and Welfare. https://eric.ed.gov/?id=ED260696

Campoy, R. (1992). The role of technology in the school reform movement. *Educational Technology, 32*(8), 17-22. http://dx.doi.org/doi/10.5555/134991.134995

Clariana, R. B., Ross, S. M., & Morrison, G. R. (1991). The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. *Educational Technology Research and Development*, *39*(2), 5-17. http://dx.doi.org/doi/10.1007/BF02298149

Clark, T. (2000). *Virtual high schools: State of the states: A study of virtual high school planning and operation in the United States*. Western Illinois University, Center for the Application of Information Technologies.

Coniam, D. (2002). Perceptions of a multimedia syllabus – making the demands of a performance test more accessible. *System 31*(1), 55-70. http://dx.doi.org/10.1016/S0346-251X(02)00073-8

Cox, M., Webb, M., Abbott, C., Blakeley, B., Beauchamp, T., & Rhodes, V. (2003). ICT and pedagogy: A review of the research literature. *ICT in Schools Research and Evaluation Series*, *18*, 41-53.

Cox, M. J., Webb, M. E., Abbott, C., Blakely, B., Beauchamp, T., & Rhodes, V. (2003). *ICT and pedagogy: A review of the research literature: A report to the DfES* (ISBN: 1844781356). https://mirandanet.ac.uk/wp-content/uploads/2019/06/ict_pedagogy.pdf

Crook, C. (1994). *Computers and the collaborative experience of learning*. Routledge.

Dexter, S. L., Anderson, R. E., & Becker, H. J. (1999). Teachers' views of computers as catalysts for changes in their teaching practice. *Journal of research on computing in education*, *31*(3), 221-239. https://doi.org/10.1080/08886504.1999.10782252

Ding, A. C. E., Ottenbreit-Leftwich, A., Lu, Y. H., & Glazewski, K. (2019). EFL teachers' pedagogical beliefs and practices with regard to using technology. *Journal of Digital Learning in Teacher Education*, *35*(1), 20-39. http://dx.doi.org/10.1080/21532974.2018.1537816

Donaldson, R., & Kötter, M. (1999). Language learning in cyberspace: Teleporting the classroom into the target culture. *CALICO Journal, 16*(4), 530-557. http://dx.doi.org/10.1558/cj.v16i4.531-557

Ducate, L., & Arnold, N. (Eds.). (2006). *Calling on CALL: From theory and research to new directions in foreign language teaching*. Computer Assisted Language Instruction Consortium.

Duffy, T., Dueber, B., & Hawley, C. (1996). Critical thinking in a distributed environment: A pedagogical base for the design of conferencing systems. In T. Liao (Ed.), *Advanced educational technology: Research issues and future potential* (pp. 51-78). Springer Verlag.

Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration? *Educational Technology Research and Development, 53*(4), 25-39. https://doi.org/10.1007/BF02504683

Ertmer, P., Addison, P., Lane, M., Ross, E., & Woods, D. (1999). Examining teachers' beliefs about the role of technology in the elementary classroom. *Journal of Research on Computing in Education, 32*(1), 54-72. https://doi.org/10.1080/08886504.1999.10782269

Ferrara, K., Brunner, H., & Whittemore, G. (1991). Interactive written discourse as an emergent register. *Written Communication, 8*(1), 8-34. https://doi.org/10.1177%2F0741088391008001002

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*(1), 59-74. https://doi.org/10.1037/h0035872

Gilakjani, A. (2017). A review of the literature on the integration of technology into the learning and teaching of English language skills. *International Journal of English Linguistics, 7*(5), 95-106. https://doi.org/10.5539/ijel.v7n5p95

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Aldine Transaction. https://psycnet.apa.org/doi/10.1002/j.2161-007X.1995.tb01014.x

Habibi, A., Abdul Razak, R., Yusop, F., Mukminin, A., & Yaqin, L. (2020). Factors affecting

ict integration during teaching practices: A multiple case study of three Indonesian universities. *Qualitative Report, 25*(5), 1127-1144. http://doi.org/10.46743/2160-3715/2020.4150

Hadley, M., & Sheingold, K. (1993). Commonalities and distinctive patterns in teachers' integration of computers. *American Journal of Education, 101*(3), 261-315. https://doi.org/10.1086/444044

Hanson-Smith, E. (2000). *Technologically enhanced learning environments*. *Case Studies in TESOL Practice Series*. Teachers of English to Speakers of Other Languages, Inc.

Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Education Technology Research and Development, 55*(3)*,* 223-252. https://doi.org/10.1007/s11423-006-9022-5

Hignite, M. A., & Echternacht, L. J. (1992). Assessment of the relationships between the computer attitudes and computer literacy levels of prospective educators. *Journal of Research on Computers in Education, 24*(3), 381-391. https://www.learntechlib.org/p/145685/

Hochman, A., Maurer, M., & Roebuck, D. (1993). Buttons and cards and fields, oh my! *Tech Trends, 38*(2), 25-28. https://doi.org/10.1007/BF02763856

Howley, A., Wood, L., & Hough, B. (2011). Rural elementary school teachers' technology integration. *Journal of Research in Rural Education, 26*(9), 1-13. http://jrre.psu.edu/ articles/26-9.pdf.

Inan, F. A., & Lowther, D. L. (2010). Factors affecting technology integration in K-12 classrooms: A path model. *Educational Technology Research and Development, 58*(2), 137-154. http://dx.doi.org/10.1007/s11423-009-9132-y

Jimoyiannis, A. (2010). Designing and implementing an integrated technological pedagogical science knowledge framework for science teachers professional development. *Computers & Education, 55*(3), 1259-1269. http://dx.doi.org/10.1016/j.compedu.2010.05.022

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). *The 2010 horizon report*. The New Media Consortium.

Jonassen, D., Peck, K., & Wilson, B. (1999). *Learning with technology: A constructivist perspective*. Prentice-Hall.

Karimi, L., Khodabandelou, R., Ehasani, M., & Ahmad, M. (2014). Applying the uses and gratifications theory to compare higher education students' motivation for using social networking sites: Experiences from Iran, Malaysia, United Kingdom, and South Africa. *Contemporary Educational Technology, 5*(1), 53-72. http://dx.doi.org/10.30935/cedtech/6115

Kearns, J. (1992). Does computer coursework transfer into teaching practice? *Journal of Computing in Teacher Education, 8*(4), 29-34. https://doi.org/10.1080/10402454.1992.10784181

Kelm, O. R. (1992). The use of synchronous computer networks in second language instruction: A preliminary report. *Foreign Language Annals*, *25*(5), 441-454. http://dx.doi.org/10.1111/j.1944-9720.1992.tb01127.x

Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *Modern Language Journal, 79*(4), 457-476. https://doi.org/10.2307/329999

Kern, R. (1996). Computer-mediated communication: Using e-mail exchanges to explore personal histories in two cultures. In M. Warschauer (Ed.), *Telecollaboration in foreign language learning* (pp. 105-120). University of Hawaii.

Kim, H. K. (2008). Beyond motivation: ESL/EFL teachers' perceptions of the role of computers. *Calio Journal, 25*(2) 241-259.

http://www.jstor.org/stable/calicojournal.25.2.241

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary issues in technology and teacher education*, *9*(1), 60-70. https://doi.org/10.1177%2F002205741319300303

Kollias, V., Mamalougos, N., Vamvakoussi, X., Lakkala, M., & Vosniadou, S. (2005). Teachers' attitudes to and beliefs about web-based collaborative learning environments in the context of international implementation. *Computers & Education*, *45*(3), 295-315. http://dx.doi.org/10.1016/j.compedu.2005.04.012

Kronnenberg, N. (1994). Developing communicative and thinking skills via electronic mail. *TESOL Journal, 4*(2), 24-27. https://www.learntechlib.org/p/78350/

Lak, M., Soleimani, H., & Parvaneh, F. (2017). The effect of teacher-centeredness method vs. learner-centeredness method on reading comprehension among Iranian EFL learners. *Journal of Advances in English Language Teaching, 5*(1), 1-10. https://european-science.com/jaelt/article/download/4886/2415

Lee, L. (1998). Going beyond classroom learning: Acquiring cultural knowledge via on-line newspapers and intercultural exchanges via online chat rooms. *CALICO Journal, 16*(2), 101-120. http://www.jstor.org/stable/24147651

Lee, L. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the US. *Language Learning & Technology, 8*(1), 83-100. http://dx.doi.org/10125/25231

Levin, T., & Wadmany, R. (2008). Teachers' views on factors affecting the effective integration of information technology in the classroom: Developmental scenery. *Journal of Technology and Teacher Education, 16*(2), 233-263. https://www.learntechlib.org/primary/p/22950/.

Mahmoudzadeh, S. (2014). The effect of using PowerPoint on Iranian EFL learners' knowledge of abstract vocabulary. *Procedia - Social and Behavioral Sciences, 98,* 1077-1084. http://dx.doi.org/10.1016/j.sbspro.2014.03.519

Marcinkiewicz, H. R. (1993). Computers and teachers: Factors influencing computer use in the classroom. *Journal of Research on Computing in Education, 26*(2)*,* 220-237. https://doi.org/10.1080/08886504.1993.10782088

Miller, L., & Olson, J. (1994). Putting the computer in its place: A study of teaching with technology. *Journal of Curriculum Studies, 26*(2), 121-141. https://doi.org/10.1080/0022027940260201

Office of Technology Assessment (1995). *Information technology and its impact on American education*. U.S. Government Printing Office.

Pajares, M. F. (1992). Teacher's beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*, 307-322. https://doi.org/10.3102%2F00346543062003307

Pedersen, S., & Liu, M. (2003). Teachers' beliefs about issues in the implementation of a student-centered learning environment. *Education Technology Research & Development, 51*(2), 57-76. http://dx.doi.org/10.1007/BF02504526

Project Tomorrow. (2011). The new 3 E's of education: Enabled, engaged, empowered: How today's educators are advancing a new vision for teaching and learning. http://www.tomorrow.org/speakup/ pdfs/SU10_3EofEducation_Educators.pdf

Rahimi, M., & Yadollahi, S. (2011). ICT use in EFL classes: A focus on EFL teachers' characteristics. *World Journal of English Language, 1*(2), 17-29. https://doi.org/10.5430/wjel.v1n2p17

Raygan, A., & Moradkhani, S. (2020). Factors influencing technology integration in an EFL context: Investigating EFL teachers' attitudes, TPACK level, and educational

climate. *Computer Assisted Language Learning (online).* https://doi.org/10.1080/09588221.2020.1839106

Riasati, M. J., Allahyar, N., & Tan, K. E. (2012). Technology in language education: Benefits and barriers. *Journal of Education and Practice, 3*(5), 25-30. https://www.iiste.org/Journals/index.php/JEP/article/view/1495/1427

Ritchie, D., & Wiburg, K. (1994). Educational variables influencing technology integration. *Journal of Technology and Teacher Education, 2*(2), 143-153. https://www.researchgate.net/publication/234796304_Educational_variables_influencing_technology_integration

Ritzhaupt, A. D., Dawson, K., & Cavanaugh, C. (2012). An investigation of factors influencing student use of technology in K-12 classrooms using path analysis. *Journal of Educational Computing Research, 46*(3), 229-254. https://doi.org/10.2190%2FEC.46.3.b

Roblyer, M. D. (2006). *Integrating educational technology into teaching* (4th ed.). Pearson Merrill Prentice Hall

Salaberry, M. R. (2000). Pedagogical design of computer-mediated communication tasks: Learning objectives and technological capabilities. *Modern Language Journal, 84*(1), 28-37. https://doi.org/10.1111/0026-7902.00050

Scherer, R., Tondeur, J., Siddiq, F., & Baran, E. (2018). The importance of attitudes toward technology for pre-service teachers' technological, pedagogical, and content knowledge: Comparing structural equation modeling approaches. *Computers in Human Behavior*, *80*, 67-80. http://dx.doi.org/10.1016/j.chb.2017.11.003

Sepehr, H., & Harris, D. (1995). Teachers' use of software for pupils with specific learning difficulties. *Journal of Computer Assisted Learning, 11*(2), 64-71. https://doi.org/10.1111/j.1365-2729.1995.tb00118.x

Soleimani, H., & Arabloo, P. (2018). Technology training courses in Iran and their effectiveness in helping teachers to use technology in their classrooms: A study. *International Journal of English Language & Translation Studies, 6*(2), 14-19.

Strommen, E., & Lincoln, B. (1992). Constructivism, technology, and the future of classroom learning. *Education and Urban Society, 24*(4), 466-476. https://doi.org/10.1177%2F0013124592024004004

Taylor, R. (1980). *The computer in the school: Tutor, tool, tutee*. Teachers College Press.

Todd, N. (1993). A curriculum model for integrating technology in teacher education courses. *Journal of Computing in Teacher Education, 9*(3), 5-11. https://doi.org/10.1080/10402454.1993.10784201

Todman, J., & Dick, G. (1993). Primary children's and teachers' attitudes toward computers. *Computers in Education, 20*(2), 199-203. https://doi.org/10.1016/0360-1315(93)90088-Z

Tondeur, J., Valcke, M., & Van Braak, J. (2008). A multidimensional approach to determinants of computer use in primary education: Teacher and school characteristics. *Journal of Computer Assisted Learning, 24*(6), 494-506. http://dx.doi.org/10.1111/j.1365-2729.2008.00285.x

Tondeur, J., Van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2016). Erratum to: Understanding the relationship between teachers? Pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development, 65*(3), 577-584. https://psycnet.apa.org/doi/10.1007/s11423-016-9481-2

Trochim, W., & Donnelly, J. (2006). *The research methods knowledge base* (3rd ed.). Atomic Dog Publishing.

Ushioda, E. (2000). Tandem language learning via e-mail: From motivation to autonomy. *ReCALL, 12*(1), 121-128. https://doi.org/10.1017/S0958344000000124

Usluel, Y. K., Mumcu, F. K., & Demiraslan, Y. (2007). ICT in the learning-teaching process: Teachers' views on the integration and obstacles. *Hacettepe University Journal of Education*, *32*, 164-178.

Violato, C., Mariniz, A., & Hunter, W. (1989). A confirmatory analysis of a four-factor model of attitudes toward computers: A study of preservice teachers. *Journal of Research on Computing in Education, 22*(1), 199-213. https://doi.org/10.1080/08886504.1989.10781915

Vygotsky, L. (1986). *Thought and language*. MIT Press.

Warschauer, M. (1995). *Virtual connections: Online activities and projects for networking language learners*. University of Hawaii, Second Language Teaching and Curriculum Center.

Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3-20). Logos International. DOI: 10.12691/jll-4-1-2

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching, 31*(1), 57-71. http://dx.doi.org/10.1017/S0261444800012970

Wetzel, K. (1993). Models for achieving computer competencies in preservice education. *Journal of Computing in Teacher Education, 9*(4), 4-6. https://doi.org/10.1080/10402454.1993.10784208

Woodrow, J. E. (1990). Locus of control and student-teacher computer attitudes. *Computers in Education, 14*(5), 421-432. https://doi.org/10.2190%2FCBG3-X2UE-DQGY-YWH9

Woodrow, J. E. (1992). The influence of programming training on the computer literacy and attitudes of preservice teachers. *Journal of Research on Computing in Education, 25*(2), 200-218. https://doi.org/10.1080/08886504.1992.10782044

Zohrabi, M., Torabi, M. A., & Baybourdiani, P. (2012). Teacher-centered and/or student-centered learning: English language in Iran. *English Language and Literature Studies, 2*(3), 18-30. https://doi.org/10.5539/ells.v2n3p18

Appendix A
Kim (2008). Beyond Motivation: EFL-ESL Teachers' Perceptions of the Role of Computers

Cleveland State University

General Information

1. Would you tell me about your computer skills? Could you name the computer software programs that you can use?
2. Have you taken any courses related to computers? If any, what are the names of the courses?

Experience of Teaching

3. Would you tell about your ESL/EFL teaching experience briefly?

Experience of Using Computers

4. Have you used computers in your teaching? How did you use computers in your teaching, if you have?

Using Computers in Language Teaching

5. What do you think of the roles of computers in language teaching?
6. What are the benefits of using computers in language teaching?
7. Are you planning to use computers in your teaching in the future? If yes, how are you going to use computers in your classroom?
8. Can you give me some examples of how you are going to use computers in your teaching?

Teacher Beliefs

9. Do you think technology-enhanced learning can better prepare students for 21st century? Explain your viewpoint?
10. What is motivating you to use technology the amount you do?
11. What is motivating you to use the math software the amount you do?

Teacher Perception

12. Do you know how much usage the product provider recommends? How much do you think the product should be used per week?
13. How much professional development did you receive related to this software?
14. How are you using the software?
15. How do you know when to use the software?

Instructional Strategies

16. What are some ways you have helped students view technology as a learning tool?
17. How do you incorporate technology into your instruction?
18. What did you do to incorporate software into your instruction?
19. What are some ways that you have been able to use technology to enhance student learning?
20. Do you encourage collaboration and creative thinking? If so how?

Motivation Strategies

21. How are students using the software?
22. How often do you check student progress?
23. Do students use this software at home?
24. What strategies do you use to motivate and encourage students?
25. What strategies do you use to keep students engaged?

Change Management

26. In your school or classroom, what strategies were used to manage this change to your instruction?
27. What instructional or management strategies do you use to integrate technology in your classroom?

Iranian teachers

28. What barriers do Iranian Teachers face up with using computers?
29. How do Iranian teachers relate course and material to computer?
30. What strategies do Iranian teachers use to homogenize the level of computers' knowledge of students?
31. Which skill is greatly promoted through computers?
32. Do Iranian teachers use the computer in their teaching as an important course or as supplementary material?

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

Locating the Source of Number Agreement Errors in Intermediate-to-Advanced Second Language Learners of English

Research Article
pp. 137-160

Ebrahim Safaie[1]

Received: 2021/02/01 Accepted: 2021/03/06

Abstract
This study explores the extent to which adult Second Language (L2) Learners of English are sensitive to Subject-Verb (S-V) agreement errors with thematic verbs and copulas. A group of intermediate-to-advanced Persian-English second language learners and a group of native English speakers (NSs) participated in an online Speeded Acceptability Judgment task. The results are as follows. Whereas NSs are sensitive to number agreement errors in both verb types, second language learners are not. For the latter group, the results reveal the following systematic errors. For agreement errors in thematic verbs, there is an omission-commission asymmetry; whereas L2 learners are not sensitive to omission errors, they are sensitive to commission errors in finite forms. For copulas, there is a singular-plural asymmetry, indicating higher error rates in **plural subject NP + is* configurations than in **singular subject NP+ are* ones. Yet, proficiency seems to be a strong predictor of L2 learners' sensitivity to agreement errors. The results support the Morphological Underspecification Hypothesis (McCarthy, 2007; 2008; 2012).

keywords: feature clash errors, omission-commission asymmetry, singular-plural asymmetry, subject-verb agreement, underspecification errors

[1] Assistant Professor, Department of English Language, Lamerd Branch, Islamic Azad University, Lamerd, IRAN. es.safaie@iaulamerd.ac.ir; safaie124@yahoo.com

DOI: 10.22051/LGHOR.2021.34969.1438

Introduction

It has often been a common observation that Subject-Verb (S-V) agreement poses a protracted difficulty for adult Second Language Learners (L2ers). This has robustly been reported in both Second Language Acquisition (SLA) (Eubank & Grace, 1998; Prévost & White, 2000; Herschensohn, 2001) and L2 processing research (Chen, Shu, Liu, Zhao, & Li, 2007; Ojima et al. 2005; Jiang, 2004; McDonald, 2006; Sato & Felser, 2010). Importantly, growing evidence from the study of S-V agreement processing during L2 listening and reading has shown that L2ers' difficulty with agreement may not be a purely production-specific problem because they are also less sensitive to S-V agreement than native-speakers (NSs) are during L2 comprehension (e.g., Wakabayashi, 1997; McDonald, 2006; Chen et al., 2007; Jiang, 2004; Safaie, 2015; Sato & Felser, 2010), as in (1) and (2).

(1) The price of the car was/*were too high. (Chen, et al., 2007, p. 163)

(2) I hear that Tom goes/*go to the pub every night. (Wakabayashi, 1997, p. 160)

Two missing gaps are observable in L2 processing research on S-V agreement compared to SLA research. First, while the majority of the aforementioned SLA studies have focused on S-V agreement with both thematic verbs and copula *be* in a single study, in L2 processing, separate research has focused on one or the other, but not on both in a single study. For instance, L2 processing studies on agreement have only tested either copula *be* (e.g., Chen et al., 2007; Jiang, 2004; Tanner, et al., 2012) or thematic verbs (e.g., Armstrong et al., 2018; McDonald, 2006; Ojima et al., 2005; Sato & Felser, 2010; Wakabayashi, 1997) ignoring the potential effects of verb type differences on L2 processing of S-V agreement. Given that a fully-fledged L2 processing account of S-V agreement should ideally include the effect of both verb types, it is surprising that no L2 processing studies to date have been specifically designed to examine both verb types in a single study so that they can develop a comprehensive perspective regarding L2 processing of S-V agreement. Second, L2 processing research has not deemed the potential effect of markedness on L2 processing of S-V agreement. Under markedness theory, marked representations are more specified than unmarked ones which are underspecified. Therefore, unmarked representations should be processed with greater difficulty than marked ones. Ignoring the potential token frequency differences in the input and word-formation processes between copulas and thematic verbs, these two verb types are different in terms of markedness as will be discussed in detail later below. To the best of the researcher's knowledge, L2 processing research of S-V agreement has not been studied from this perspective.

To fill the missing gaps in L2 processing research, this study aims at exploring L2ers' variability in S-V agreement from the perspective of markedness theory using both verb types in a single study in Persian speakers of L2 English. To this end, both verb types (thematic verbs vs. copula *be* with their possible configurations when combined with both singular and plural subjects) are examined to see how L2ers treat them online in a Speeded Acceptability Judgment (SAJ) task when they are under added processing pressure.

The current study is significant because, first, it tests both verb types in a single study. This allows a unique way of exploring S-V agreement within a single participant. Second, this study tests Persian speakers of L2 English. Persian as the L1 of the L2ers is highly inflected for both verb types. Thus, studying Persian speakers is likely to rule out the possibility that the variability might be related to an absence of similar features in the L1 as reported in some studies (e.g., Hawkins &

Chan, 1997; Franceschina, 2005). This is because features like person and number agreement have already been acquired via the L2ers' L1 before the critical period is over. Third, this is the first study applying markedness theory to the L2 processing of S-V agreement.

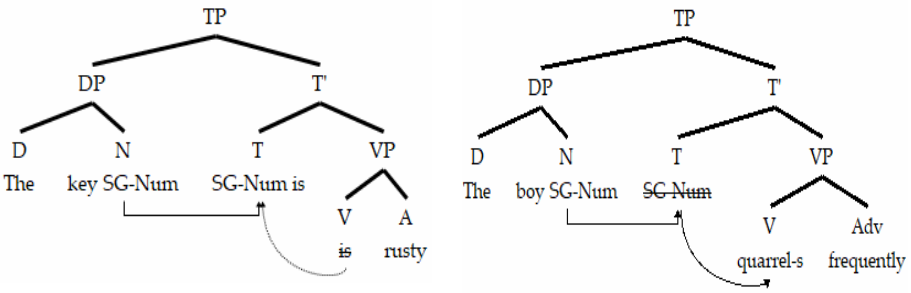
To apply the markedness theory, the present study tests the predictions of the Morphological Underspecification Hypothesis (MUH) (McCarthy, 2012; 2008) in the L2 processing of S-V agreement. As a point of departure, after presenting a minimalist account of S-V agreement, the theoretical framework of the MUH is discussed along with evidence from the L2 processing research supporting the MUH predictions.

The Minimalist Account of S-V Agreement with Thematic Verbs and Copulas

Agreement is a computational process at the syntax-morphology interface. It is syntactic as it depends on lexical categories like N, V, etc. and involves MERGE and MOVE operations. It is morphological as it affects the forms of morphemes. Syntactically, for S-V agreement with both copulas and thematic verbs, successive applications of several MERGE operations lead to the formation of the basic skeleton in the syntactic component (Franck et al., 2006), as displayed in Figure 1 (in the tree diagrams, details are omitted for simplicity). Afterwards, the unvalued person and number features [u-Pers, u-Num] of T(ense) are valued by the already valued features of the subject phrase demonstrated with solid trace lines. Then, in the Phonetic Form (PF) component, HEAD MOVEMENT rules operate in different directions depending on the type of the verb to spell out their phonetic representations. Copulas raise to T position to receive their feature values from T. But thematic verbs stay *in situ* and are inflected via affix-hopping/lowering (Radford, 2009), as illustrated with curved trace lines for both verb types.

Figure 1

Tree Diagram Displaying Raising Possibilities



a. Verb-raising **b. Affix-lowering/hopping**

It should be pointed out in passing that several SLA researchers, (Ionin & Wexler, 2002; Prévost & White, 2000) suggest that, in the initial stages of learning L2 English, more efficient use of copulas than inflected thematic verbs is due to the different raising possibilities of these two verb types. That is, L2 beginners master suppletive *be* forms which raise to T but not 3rd person singular present tense affixal *–s* which lowers to T leading to a suppletive-affixal asymmetry (i.e., L2ers' easy mastery of S-V agreement with suppletive *be* forms than that with thematic verbs). If this argument is true, L2ers at the higher stages of learning should not have problems in S-V agreement with both verb types, because they should be beyond

this initial stage. However, these types of problems persist even for advanced L2ers as will be discussed later below.

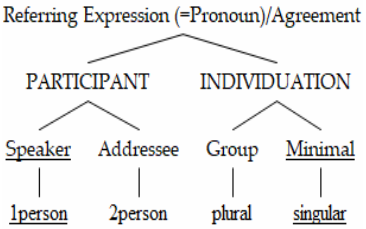
Morphological Underspecification Hypothesis

Using a recent theoretical framework in morphological theory, known as the Feature Geometric Hypothesis (Cowper, 2005; Harley, 1994; Harley & Ritter, 2002) McCarthy (2007; 2008; 2012) suggests the MUH to account for L2 variability in agreement features like *person*, *number, and finiteness*. For instance, the Feature Geometric Hypothesis posits that L1 acquirers pick up feature categories like person and number which are made up of feature values. These features gradually become hierarchically organised with privative values (as opposed to binary, e.g., [+plural], [−plural]) in the mental lexicon such that one member is marked or specified while the other is unmarked and thereby default or underspecified.

As illustrated in Figure 2 below, for Person/Number agreement, unmarked features (underlined forms) are realised as underspecified feature values, compared to more marked or specified features (not underlined). To contrast 1st and 2nd person, Addressee as the 2nd person is more marked than Speaker as the 1st person, i.e., 2nd person is marked relative to 1st person, which is unmarked. However, relative to 1st and 2nd person, 3rd person is assumed to be universally unmarked. That is why 3rd person is absent from the PARTICIPANT node. To contrast singular and plural number, INDIVIDUATION representing number, is an organising node encompassing its daughter nodes, namely Group ([plural]) which is marked, relative to the Minimal ([singular]) feature which is unmarked. Note that the marked or specified feature value is encoded by an additional node (i.e., the presence of Group and Addressee) while unmarked ones lack additional nodes (here they are underlined to demonstrate their absence from the mental lexicon).

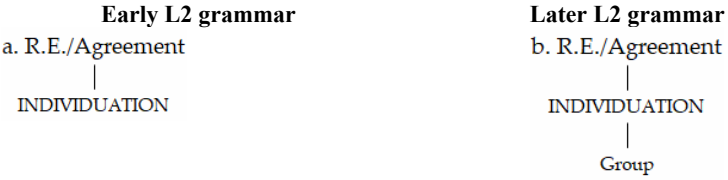
Figure 2

The Feature Geometry for S-V Agreement (extracted from Harley & Ritter, 2002)



It is postulated that marked feature values are learned later than unmarked ones because they consist of more nodes than unmarked ones. So for the developmental stages of number learning, following Harley and Ritter (2002) for the L1 grammar, McCarthy (2012, p. 34) further assumes that the early L2 grammar learns Individuation instantiating Singular in (3a) before the later L2 grammar learns Group representing plural in (3b) because Group has an additional node and more marked than Singular which is unmarked and absent from the mental lexicon.

(3)



Accordingly, under the MUH, McCarthy (2007; 2008; 2012) argues that L2 morphological errors involve a systematic substitution of unmarked (i.e., underspecified) forms as defaults across comprehension and production. For instance, McCarthy (2007) tested S-V number agreement in intermediate-to-advanced L1 English speakers of L2 Spanish (N=13) via a production task (interview). Following the Feature Geometric Hypothesis (Harley & Ritter, 2002), she found a singular-plural asymmetry in that advanced L2ers committed underspecification errors (92%) in (e.g., *Los italianos puede entender un poco "The Italians PL can 3SG understand a little (Spanish)") in which they selected singular verbs (i.e., unmarked/underspecified forms) preceded by plural subjects. This is in contrast to errors of feature clash which constituted only (8%) of L2ers' data (i.e., L2ers selected plural verbs (i.e., marked/specified forms) to substitute singular verbs preceded by the singular subjects, e.g., *he are). See more evidence for the MUH from L2ers of Spanish (López Prego & Gabriele, 2014; McCarthy, 2012) and German (Slabakova, 2009). As for the L2 processing of S-V agreement, to the best of the researcher's knowledge, there have been no studies testing the MUH. Thus, after discussing English verbs in terms of the Feature Geometric Hypothesis, evidence will be provided from the existing processing research on L2 English which seems to support the MUH.

Feature Geometric Hypothesis: Thematic Verbs vs. Copulas of English

Given the assumptions of the Feature Geometric Hypothesis, the marked-unmarked relations shown in Table 1 are applied to the verbs of English. In SUBJECT + COPULA BE (S-VC) dependencies, for person with participant as its feature category, non-3rd persons (participant) are marked but 3rd is universally unmarked. For number with individuation as its feature category, plural (Group) is marked but singular (Minimal) is universally unmarked. For FINITENESS in SUBJECT + THEMATIC VERB (S-VT) dependencies, following Cowper (2005), the non-finite form is unmarked relative to the finite form which is marked (for a detailed discussion of the range of markedness criteria, see Harley & Ritter, 2002).

Table 1

Summary of Markedness Relations in Feature-Geometric Theory

Verb type	Feature category	Marked Feature value	Unmarked Feature value
Copula be	Person (PARTICIPANT)	Non-3rd (=1st or 2nd)	3rd
	Number (INDIVIDUATION)	Plural (Group)	Singular (Minimal)
Thematic verb	FINITENESS	Finite	Nonfinite

Note: This summary was taken from McCarthy, 2012, p. 33, but applied to verbs of English with some modifications.

Drawing on the Feature Geometric Hypothesis to explain L2 morphological variability, the MUH holds that "L2 errors are instances of underspecification, not

feature clash" (McCarthy, 2008, p. 467; 2012, p. 33). For concreteness, as illustrated in Table 2, for thematic verbs, errors of underspecification refer to omission errors where non-finite unmarked forms substitute finite marked forms, as in *She speak, in obligatory contexts. Feature clash occurs where an incorrectly marked form gets inserted, as in *They speaks, when such a form is not obligatory. For copulas, [singular] is universally unmarked compared to [plural] which is marked. This means that when the subject is singular and the copula be is plural [are] this is an instance of feature clash because the syntactic context changes from an unmarked form [singular subject] to a marked form of the verb [plural copula be]. In contrast, when the subject is plural and copula be is singular [is], this is an instance of underspecification because the syntactic context changes from a marked form to an unmarked form.

Table 2
Error Types Based on Markedness Relations in Feature Geometric Theory

Verb type	Error type	Example	Predictions
Thematic verb	Underspecification Marked⇒ Unmarked	Finite ⇒ Nonfinite e.g.**She speak*	More errors with nonfinite underspecified forms (*speak*)
	Feature clash Unmarked⇒ Marked	Nonfinite ⇒Finite e.g.**They speaks*	Fewer errors with finite specified forms (*speaks*)
Copula *be*	Underspecification Marked⇒ Unmarked	Specified(Plural*NP*)⇒ Underspecified (Singular *is*) e.g. **The spoons is …*	More errors with underspecified forms (*is*)
	Feature clash Unmarked ⇒ Marked	Underspecified (Singular NP)⇒ Specified (Plural *are*)e.g. **The spoon are …*	Fewer errors with specified forms (*are*)

Evidence for the MUH from L2 Processing of English S-V Agreement

Thematic Verb

According to the top part of Table 2, L2 processing studies using thematic verbs show that L2ers exhibit different patterns depending on their proficiency levels consistent with the predictions of the MUH. Whereas intermediate L2ers are not sensitive to either error type, advanced L2ers are sensitive to errors of feature clash but not to errors of underspecification. For instance, in an Event-Related Potentials (ERP) experiment, Ojima et al., (2005) investigated ERP responses to S-V agreement violations (e.g., *Turtles moves slowly vs. Turtles move slowly) in L1 Japanese L2ers of English who were either moderately (intermediate) or highly proficient (advanced). Like the NSs who were sensitive to errors of feature clash by showing a biphasic LAN-P600 pattern, the highly proficient L2ers were sensitive to them by showing a LAN (but no P600) typically associated with syntactic processing. In contrast, the moderately proficient L2ers were not sensitive by showing none of the ERP components. (see also Rossi et al., 2006) in which the NSs and the high proficiency learners (L1 Italian learners of L2 German & L1 German learners of L2 Italian) showed a LAN and a P600 effect in response to feature clashes in S-V agreement, and Armstrong et al.'s (2018) study, in which both NSs and intermediate-to-advanced Chinese L2ers of English, showed a P600 effect in response to errors of feature clash, as in The cookies taste/*tastes …).

Likewise, behavioural studies such as Wakabayashi (1997) and Sato & Felser (2010) showed that whereas advanced L2ers were not sensitive to errors of underspecification they were sensitive to those of feature clash in S-VT

dependencies. Studying L2ers' online sensitivity to S-V agreement, Wakabayashi (1997) focused on ungrammaticality caused by overuse (i.e., errors of feature clash) and omission of 3rd person singular present tense –s (errors of underspecification) by using a Self-Paced Reading (SPR) task. NSs and two groups of L2ers (advanced & intermediate Japanese L2ers of English) were tested. The results revealed that NSs were sensitive to both error types. As for L2ers, whereas the intermediate Japanese L2ers of English (JLE) were insensitive to both errors of commission/feature clash and omission/underspecification errors in number disagreement, the advanced JLEs were insensitive only to omission/ underspecification errors. Similarly, Sato and Felser (2010) investigated sensitivity to omission errors in S-V agreement vis-à-vis case violations (S-V agreement: *He frequently yawn vs. Accusative case: *He admires she) in intermediate-to-advanced L2ers of English from three typologically different L1 backgrounds (German, Japanese, Chinese) using a SAJ task. The results showed that all L2ers, independently of their L1s, were less sensitive to S-V agreement (i.e., an underspecified form like yawn) than case violations, see also McDonald, 2006, for similar findings. Summarising, in line with the MUH, for thematic verbs, these observations may suggest that morphological underspecification is more likely to be the source of variability in L2 processing of S-V agreement and proficiency seems to modulate it.

Copula be

According to the bottom part of Table 2, (except for Jiang's (2004) study) L2 processing studies on S-V agreement with copula be, show a singular-plural asymmetry; L2ers perform better in *singular subject NP + plural copula be configurations (errors of feature clash) than in *plural subject NP+ singular copula be ones (errors of underspecification).

Using an SPR task, Jiang (2004) examined S-V agreement in NSs and a group of Chinese intermediate L2ers. In his experiments 2 and 3, he tested S-V agreement violations in four conditions, as in The bridges to the island were about… (Subject PL + Attractor SG + Verb PL (PSP)) vs. *The bridge to the island were about… (Subject SG + Attractor SG + Verb PL (SSP)) and The bridge to the island was about… (Subject SG + Attractor SG + Verb SG (SSS)) vs. *The bridges to the island was about… (Subject PL + Attractor SG + Verb SG (PSS)). The results showed significant longer RTs at the verb and the following region for the NSs. As for the L2ers, while they also showed longer RTs at the same regions (even longer than the NSs) but their results were not significant. Note, however, that Jiang's study can be criticised for a couple of reasons. First, the comparison of PSP with SSP conditions resulted in slowdowns on the verb region for both NSs and L2ers but NSs showed only significant results in participant analysis but not item analysis. So, both groups almost converged because SSP is an instance of feature clash as the syntactic context changes from an unmarked singular subject NP to a marked plural copula be form. The comparison of SSS with PSS, however, resulted in a NS-L2er contrast. Perhaps this is because the PSS condition is an instance of underspecification error (because the syntactic context changes from a marked plural subject NP to an unmarked singular copula be form) and this kind of error is committed by the NSs, too.[1] Thus, L2ers almost performed better in SSP than in PSS showing a singular-

[1] According to the CORPUS OF CONTEMPORARY AMERICAN ENGLISH (COCA) database (Davies, 2008) configurations like *$Subject\ NP_{PL}$ + *is* (e.g., **The spoons is*), show a

plural asymmetry. Second, regarding the statistical analyses used, Jiang's study is questionable because although L1 group (NSs vs. Chinese) was an independent between-group factor, the data were analysed separately for each group. That is, by using a within-condition design, separate paired t-tests were used for each condition and each group. Consequently, it is not clear whether groups were different if a mixed design had been used.

However, in contrast to Jiang's study[1] which was claimed to show L2ers'reduced sensitivity to S-V agreement in S-VC dependencies, Tanner et al. (2012) found their native-like sensitivity. They used ERPs in a study of L2 agreement processing testing native speakers and advanced L1 Spanish L2ers of English. The critical ungrammatical items (*The key to the wooden cabinet are rusty (Subject SG + Attractor SG + Verb PL (SSP)) and (*The key to the wooden cabinets are rusty (Subject SG + Attractor PL + Verb PL (SPP)) had a singular subject and a plural verb which are instances of feature clash because the plural copula are is marked and thereby clashes with the [singular] feature of the subject, hence no processing difficulty for advanced L2ers (following McCarthy, 2007; 2008; 2012). Similarly, using eye tracking, Lim and Christianson (2015) found that both NSs and higher proficiency Korean-English L2ers were sensitive to agreement violations (The teachers who instructed the student were very strict (PSP) vs. *The teacher who instructed the student were very strict (SSP). Note that the ungrammatical critical items had singular subjects and plural copula verbs, hence these are instances of feature clash. Altogether, it can be concluded that in these studies a singular-plural asymmetry was observed; L2ers were better in singular subject NP + plural copula be configurations (errors of feature clash) than in plural subject NP+ singular copula be ones (errors of underspecification).

Summarising, the results of L2 processing studies of S-V agreement in thematic verbs and copulas can provide compelling evidence for the MUH, hence morphological underspecification is the culprit for variability in S-V agreement. Additionally, the L2 processing studies revealed the paramount role of L2ers' proficiency. Whereas moderate L2ers were not sensitive to either violation type, advanced L2ers were sensitive to marked feature values (i.e., errors of feature clash) but not to unmarked feature ones (i.e., errors of underspecification).

This Study

Building upon the MUH (McCarthy, 2007; 2008; 2012), the present study explores sensitivity to S-V agreement errors (underspecification & feature clash errors) in the intermediate-to-advanced L2ers and NSs of English using different verb types (copulas vs. thematic verbs). This was investigated by measuring the L2ers' reactions to errors (ungrammatical forms) in an online SAJ task. The focus is on errors because the existing research shows that L2ers might not have problems with correct forms of S-V agreement in grammatical items for which verbs are inflected accurately (Lardiere, 2007; McCarthy, 2012; 2008; Prévost & White,

---

co-occurrence frequency of 0.06 (per million) words in the COCA corpus (1990-2012). Moreover, the British National Corpus shows a co-occurrence frequency of 0.20 (per million) for **Subject NP* $_{PL}$ + *is*.

[1] Chen et al. (2007) also tested Chinese L2ers using ERP components. However, as the researchers themselves pointed out, their study did not "examine the S-V agreement errors actually made by L2 learners" (p. 171), rather they studied agreement attraction for which NSs have been documented to show errors (see e.g., Pearlmutter, et al. 1999).

2000). Moreover, the SAJ task was used because processing pressure seems to be a potential factor giving rise to L2 variability (Ionin & Wexler, 2002; Prévost & White, 2000). According to the existing L2 processing research and the MUH discussed above, the following questions and predictions are explored:

Question 1: Are L2ers more sensitive to errors of feature clash than underspecification?

Prediction 1: According to the MUH, "L2 errors are instances of underspecification, not feature clash" (McCarthy, 2008, p. 467; 2012, p. 33) (cf. Table 2).

Question 2: Does proficiency modulate their sensitivity to these two error types?

Prediction 2: Proficiency is a key factor in their sensitivity to error types.

Method

Instrument

Speeded Acceptability Judgment (SAJ) Task. A SAJ task requires that participants react to the acceptability of stimuli as quickly and accurately as possible. Stimuli are broken into words or phrases and presented at a very high-speed rate. This task uses a Rapid Serial Visual Presentation (RSVP) mode which presents one word at a time for an exactly short period.

Insensitivity leads to higher error rates in ungrammatical items relative to sensitivity which results in lower error rates. (McElree & Griffith, 1995). This task is useful for studies exploring implicit processing (Schütze, 1996; Ellis, 2005). This is because the added time pressure in this task may make the effect of explicit knowledge or metalinguistic judgments less likely to influence L2 response behaviour. Additionally, an SAJ task is appropriate to see whether processing pressure is likely to cause L2 variability

Participants

Thirty-two L1 Persian speakers of L2 English (19 males; 13 females; 1 left-handed) were recruited from among the Iranian student community at the University of Essex and Manchester. All participants had a normal or corrected-to-normal vision. All non-native participants were residents in the UK at the time of testing and, on average, had been exposed to British English for a mean of 2.39 in years. The bio-data and their mean proficiency scores as measured by the Oxford Placement Test (OPT) (Allan, 1992) are given in Table 3. The learner group scored above 65 which corresponds to the intermediate-to-advanced level on the OPT scale. This is the minimum requirement for L2ers' participation in this study.

Table 3

L2ers' Bio-Data and OPT

	Mean	SD	Range
Age (years)	27.28	3.5	22-34
Age of onset (Age of first exposure to English in years)	11.28	2.29	6-17
Length of residence (years)	2.39	1.28	0.6-5
OPT (total 100)	79.18	7.47	65-93

OPT: Oxford Placement Test

This experiment was also administered to a control group of thirty-three NSs of English (mean age: 23.54, age range: 18-33, 21 female, 12 male) recruited from among the students of the University of Essex, who were paid a small fee for their participation.

Materials

In the SAJ task, participants made acceptability judgments for 106 sentences, each seven words long. Thirty-two were experimental items relevant for testing thematic verbs (N = 16) vs. copula *be* (N = 16), 16 for testing simple past tense verbs, and 58 were fillers. Out of 32 experimental items, 64 sets of 8 conditions, as in (4) and (5), were distributed in 4 lists based on a Latin Square design for each verb type totaling 128 experimental items.

Items for the thematic verb, as in (4), were made by the researcher and assessed by NSs to see if they were acceptable in English. They were twenty sentences, which were modified and reduced to 16 after NSs' recommendations. Items for the copula *be* condition, as in (5), were a modified version of Pearlmutter et al.'s (1999).[1]

(4)	T-SSS G	The boy in the house quarrels frequently.	Thematic Verbs
	T-SSP U	* The boy in the house quarrel frequently.	
	T-PSP G	The boys in the house quarrel frequently.	
	T-PSS U	* The boys in the house quarrels frequently.	
(5)	C-SSS G	The slogan on the poster is expressive.	Copula *Be*
	C-SSP U	*The slogan on the poster are expressive.	
	C-PSP G	The slogans on the poster are expressive.	
	C-PSS U	*The slogans on the poster is expressive.	

T = Thematic; **C** = Copula *be*; **S** = Singular; **P** = Plural; **G** = Grammatical; **U** = Ungrammatical

Each sentence consisted of a singular or a plural head NP (e.g., the boy/s) followed by a modifier PP consisting of a preposition (e.g., in) and a singular local NP (e.g., the house) which was the object of the preposition. The subject NP part of the sentence was followed by a simple present tense form of copula be or a thematic verb and an adjective or an adverb. The function of the PP modifier was to break the co-occurrence effect between adjacent subjects and verbs which might affect L2 processing of S-V agreement. For instance, according to the CORPUS OF CONTEMPORARY AMERICAN ENGLISH (COCA) database (Davies, 2008), regardless of pronoun + copula be configurations which really add more frequencies to the use of copulas, even the Subject NP SG + is configuration (e.g., the student is) alone shows a co-occurrence frequency of 12.23 (per million) words in English.

Experimental items appeared in eight different conditions that were created by manipulating grammaticality (grammatical vs. ungrammatical) and the number of the head noun (singular vs. plural), and verb type (Thematic verb vs. Copulas). Mismatching the head noun number and the verb number created ungrammatical versions. The 32 experimental items testing S-V agreement were arranged in four lists using a Latin Square Design such that each participant saw exactly one version of each item in only one condition. Past tense items and fillers each were arranged in two lists using a Latin Square Design. The experimental items (N = 32) with four lists were combined with two lists of the past tense condition (N = 16) and two lists of the fillers (N = 58) totaling 106 items. Subsequently, the order of the items was pseudo-randomised for each presentation list such that items of the same condition did not appear consecutively.

[1] The full list of experimental items is available upon request.

Procedure

The experiment was administered individually in a soundproof room. The participant groups (the NSs and the L2ers) were first asked to fill in a questionnaire providing their bio-data and a consent form of participation. After reading the instruction on the monitor, the participants were also given the chance to raise questions regarding the procedure, if necessary. The participants were required to judge the acceptability of the sentences presented to them on a 14-inch computer screen one word at a time. Each sentence trial started with a fixation cross displayed in the centre of the screen for 500ms. Afterwards, a sentence was presented in the centre of the computer screen in a word-by-word fashion at a rate of 400ms per word (Lago & Felser, 2018). By using the RSVP paradigm, the DMDX software (Forster & Forster, 2003) presented the words automatically replacing one another. All words appeared in white letters on a black background in Arial Font of 30 point. The participants were required to make acceptability judgments after reading the final word as quickly and as accurately as possible. They used a game-pad by pressing either 'Yes' or 'No' buttons to accept or reject the sentences. The game-pad was activated from the onset of the final word. Three breaks were offered throughout the experiment after each set of items. The experiment began with seven practice items giving no clues about its critical items. It took approximately 25-35 minutes. The L2ers were also given an OPT which almost took 50-60 minutes.

Analysis

For the statistical analysis, the data set was analysed using Generalised Linear Mixed Models (GLMMs) with a logistic link function and binomial variance (e.g., error rates = accurate vs. inaccurate) (Baayen, 2008; Jaeger, 2008). GLMMs were applied using R (R Development Core Team, 2017).

Following Barr, Levy, Scheepers, and Tily (2013), the initial full models consisted of all fixed factors and a maximal random structure. When any individual random variable reached a high correlation of +l or –l, it was removed from the maximal random structure. Fixed-effects were compared through contrasts between levels. Each level of a factor was contrasted to a specified reference level shown in **bold type** below. The initial full model consisted of Group (**NSs** vs. L2ers**)**, Grammaticality (**GRAMM** vs. UNGRAMM), Verb Type (**COP** vs. THEMA**),** and Head Noun (**SG** vs. PL**)**.

To avoid issues of collinearity, predictor variables were grand-mean centred. Models were first fitted to the full data set for both groups; any interaction terms were further explored by analysing data from each group separately. The model fitted to the data from the L2ers also tested whether proficiency was a significant predictor of their error rates.

For further analyses, basic decision trees were also used. Basic decision tree models are either regression tree models, appropriate to numeric response variables or conditional inference recursive classification/partitioning trees (henceforth, classification trees for short), appropriate to binary response variables. Classification trees are a simple non-parametric regression approach (drawn via ctree () function from party package in R) that partitions the data into subsets called nodes or partitions in a data-driven way (Hothorn, Hornik, & Zeileis, 2006). Their main characteristics are classifications. They make predictor variables split recursively into a set of nodes, based on the binary response values, e.g., errors (ACCURATE = 0 vs. INACCURATE = 1) in the current study. The partitions are made

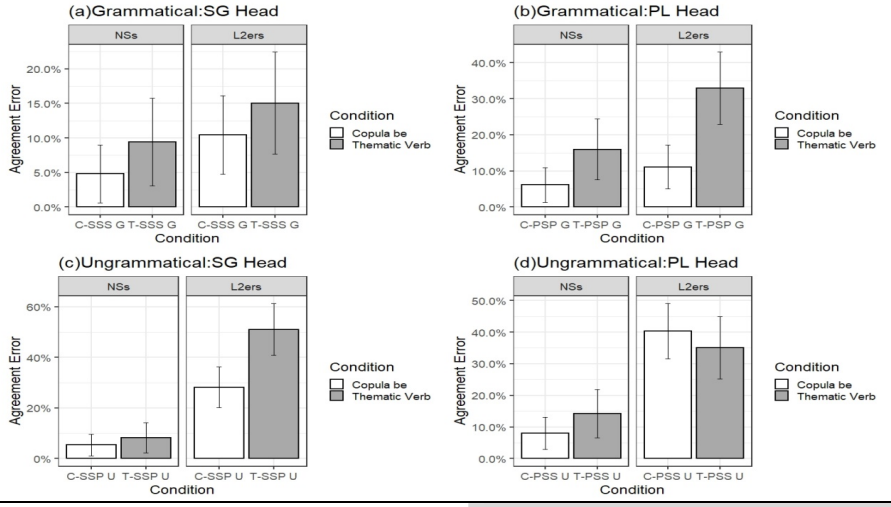
such that observations with similar error response values are grouped together. For each partition/node a constant value of error rates is predicted within that node. To partition the full data set, the algorithm starts with the root node at the top. The root node represents the full data set. The algorithm works through all predictors dividing the data into subsets where appropriate and then recursively looks into each of the subsets until further partitioning is not justified. The algorithm splits the input data into subsets that are increasingly homogeneous with the levels of the response variable. Classification trees provide more flexibility than linear regression models for modelling interactions (Blom & Baayen, 2012). Accordingly, following Blom and Baayen, (2012), along with the regression models, the current study also uses classification trees to supplement regression models in a data-driven way.

Results

The results will be presented along with figures followed by tables. Tables list the estimated coefficients, the standard errors, Z-values along with their associated p-values for the fixed effects from the models. Figure 3 illustrates the descriptive bar graph for the mean percentage of S-V agreement errors.

Figure 3

S-V Agreement with Copulas vs. Thematic Verb for both Groups (Error ars are 95% CI).



	Copula *be*		Thematic verb	
GRAMM	C-SSS G	The slogan on the poster is expressive.	T-SSS G	The boy in the house quarrels frequently.
	C-PSP G	The slogans on the poster are expressive.	T-PSP G	The boys in the house quarrel frequently.
UNGRAMM	C-SSP U	*The slogan on the poster are expressive.	T-SSP U	*The boy in the house quarrel frequently.
	C-PSS U	*The slogans on the poster is expressive.	T-PSS U	*The boys in the house quarrels frequently.

In Table 4, the coefficients for the main effects of Group, Grammaticality, and Verb Type were significant. These significant main effects indicate that the L2ers made more errors than the NSs, error rates increased in ungrammatical items relative to grammatical ones, and that error rates increased in thematic verbs compared to copulas.

The two-way interaction (Group× Grammaticality) indicates that, relative to the NSs, the L2ers displayed higher error rates in ungrammatical items than grammatical ones. However, this interaction is superseded by the four-way interaction (Group × Grammaticality × Head Noun × Verb Type) demonstrating between-verb type differences. The negative coefficient for this latter interaction reflects the fact that error rates slightly decrease in thematic verbs relative to copula be in the L2ers when the head noun is plural and the sentence is ungrammatical (cf. Panel d of Figure 3). This is in contrast to other combinations of head noun and verb type in which copula be is associated with low error rates (cf. panels a, b, & c).

Table 4

Fixed-effects from generalised linear mixed model fit to data from NSs & L2ers.

Fixed Effects	Estimate	Std. Error	z value	Pr(> |z|)	
(Intercept)	-1.97488	0.13901	-14.207	< 2e-16	***
Group (L2ers)	1.5879	0.18291	8.682	< 2e-16	***
Grammaticality (Ungrammatical)	0.52558	0.24648	2.132	0.03298	*
Head Noun (PL)	0.2782	0.21429	1.298	0.19421	
Verb Type (THEMA)	0.75172	0.25679	2.927	0.00342	**
Group (L2ers)×Grammaticality(Ungrammatical)	1.56607	0.36748	4.262	2.03e-05	***
Group (L2ers)× Head Noun (PL)	-0.05052	0.36177	-0.14	0.88893	
Grammaticality (Ungrammatical) × Head Noun (PL)	-0.55496	0.47715	-1.163	0.24479	
Group (L2ers) × Verb Type (THEMA)	-0.25282	0.35648	-0.709	0.47819	
Grammaticality (Ungrammatical) × Verb Type (THEMA)	-0.48588	0.45854	-1.06	0.28931	
Head Noun (PL) ×Verb Type (THEMA)	0.1474	0.44874	0.328	0.74254	
Group (L2ers) ×Grammaticality (Ungrammatical) × Head Noun (PL)	-0.2522	0.73294	-0.344	0.73078	
Group (L2ers)× Grammaticality (Ungrammatical) × Verb Type (THEMA)	-0.45914	0.70735	-0.649	0.51628	
Group (L2ers)× Head Noun (PL) × Verb Type (THEMA)	-0.75603	0.70832	-1.067	0.28581	
Grammaticality (Ungrammatical) × Head Noun (PL) × Verb Type (THEMA)	-0.88689	0.99346	-0.893	0.37201	
Group (L2ers)× Grammaticality (Ungrammatical) × Head Noun (PL) × Verb Type (THEMA)	-3.13841	1.42522	-2.202	0.02766	*

Formula in R: Error Rate ~ 1 + Group * Grammaticality * Head Noun * Verb Type + (1|Item) + (1+ Grammaticality * Head Noun * Verb Type |Participant)

The four-way interaction is split by Group for further analysis in Table 5 for NSs and L2ers separately. The analysis in Table 5 revealed that whereas the NSs showed no significant effects in any factors, the L2ers exhibited a different pattern. The negative coefficient for the significant main effect of proficiency shows that as proficiency increased the L2ers' error rates decreased, overall. The positive coefficient for the significant main effect of Grammaticality shows that the L2ers' error rates increased in ungrammatical items relative to grammatical ones indicating less sensitivity to ungrammatical verbs. The positive coefficient for the significant main effect of Verb Type shows that the L2ers' error rates increased in thematic verbs compared to copulas. The three-way interaction (Grammaticality × Verb Type × Head Noun) with the negative coefficient indicates that whereas errors were roughly higher in thematic verbs relative to copula be, in plural head ungrammatical sentences (panel d of Figure 3) errors were lower in thematic verbs compared to copula be. However, the significant four-way interaction (Proficiency × Grammaticality ×Verb Type× Head Noun) with the positive coefficient indicates that as proficiency increased error rates increased in thematic verbs relative to copula be in plural conditions with ungrammatical items (panel d). This means that proficiency has changed the effect of the three-way interaction such that higher proficiency L2ers detected copula be with lower error rates than thematic verbs in all four panels including panel d. In contrast, the less proficient L2ers were better in copula be than thematic verbs in all conditions except in panel d.

Table 5
Fixed-Effects from Generalised Linear Mixed Model Fit to NSs & L2ers Data Separately

NSs (Grammatical vs. Ungrammatical)					
Fixed effects	Estimate	Std. Error	z value	Pr(> |z|)	
(Intercept)	-3.496	0.4167	-8.39	<2e-16	***
Grammaticality (Ungrammatical)	-0.3311	0.7579	-0.437	0.662	
Verb Type (THEMA)	0.996	0.7139	1.395	0.163	
Head Noun (PL)	0.4941	0.7222	0.684	0.494	
Grammaticality (Ungrammatical) × Verb Type (THEMA)	-0.1725	1.4875	-0.116	0.908	
Grammaticality (Ungrammatical) × Head Noun (PL)	0.1567	1.4845	0.106	0.916	
Verb Type (THEMA) × Head Noun (PL)	0.7199	1.3468	0.535	0.593	
Grammaticality (Ungrammatical) × Verb Type (THEMA) × Head Noun (PL)	-0.1382	2.9571	-0.047	0.963	
Formula in R: ErrorRate~1 + Grammaticality * Head Noun * Verb Type+(1 |Item)+(1+ Head Noun * Grammaticality * Verb Type |Participant)					

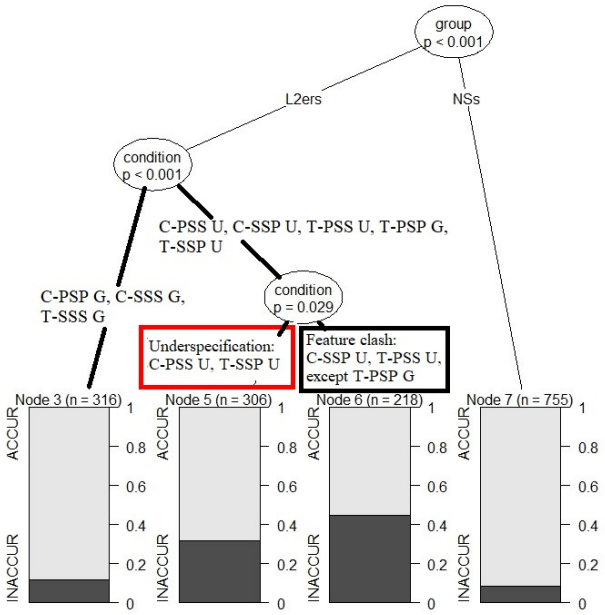
L2ers (Grammatical vs. Ungrammatical)					
Fixed effects	Estimate	Std. Error	z value	Pr(> |z|)	
(Intercept)	-1.32106	0.137057	-9.639	< 2e-16	***
Proficiency	-0.06749	0.015507	-4.352	1.35e-05	***
Grammaticality (Ungrammatical)	1.474526	0.327932	4.496	6.91e-06	***
Verb Type (THEMA)	0.725903	0.268478	2.704	0.0068	**
Head Noun (PL)	0.168137	0.270602	0.621	0.5343	
Proficiency × Grammaticality (Ungrammatical)	-0.05358	0.04052	-1.322	0.1860	
Proficiency × Verb Type (THEMA)	0.005477	0.030582	0.179	0.8578	
Grammaticality (Ungrammatical) × Verb Type (THEMA)	-0.76416	0.578925	-1.32	0.1868	
Proficiency × Head Noun (PL)	0.040031	0.032385	1.236	0.2164	
Grammaticality (Ungrammatical) × Head Noun (PL)	-0.4895	0.55575	-0.881	0.3784	
Verb Type (THEMA) × Head Noun (PL)	-0.1498	0.535856	-0.28	0.7798	
Proficiency × Grammaticality (Ungrammatical) × Verb Type (THEMA)	-0.04363	0.069573	-0.627	0.5306	
Proficiency ×Grammaticality (Ungrammatical) × Head Noun (PL)	-0.0441	0.068015	-0.648	0.5167	
Proficiency × Verb Type (THEMA) × Head Noun (PL)	0.049604	0.065057	0.762	0.4457	
Grammaticality (Ungrammatical) × Verb Type (THEMA) × Head Noun (PL)	-2.56741	1.149341	-2.234	0.0255	*
Proficiency × Grammaticality (Ungrammatical) × Verb Type (THEMA) × Head Noun (PL)	0.280833	0.141751	1.981	0.0475	*
Formula in R: ErrorRate~1 + Proficiency * Grammaticality * Head Noun * Verb Type+(1 |Item)+(1+ Head Noun * Grammaticality * Verb Type |Participant)					

To locate the loci of S-V agreement errors in participant groups, in a data-driven way, a binary recursive partitioning (the classification tree in Figure 4) was run to see which pairs among the 8 conditions ("C-PSS U" "C-PSP G" "C-SSP U"

"C-SSS G" "T-PSS U" "T-PSP G" "T-SSP U" "T-SSS G") are difficult for L2ers. Figure 4 depicts interactions between Group and Condition. Starting at the top of the tree, which represents the full data set, we observe that a first partition was made based on Group. The NSs were different from the L2ers in that they showed no difficulty in any conditions (around less than 10% error). Within the L2ers' data, a second partition was made based on Condition. They reacted differentially to ungrammatical conditions (C-PSS U, C-SSP U, T-PSS U, T-SSP U), except for T-PSP G, and grammatical conditions (C-PSP G, C-SSS G & T-SSS G). They were less accurate in the former than in the latter. Within the ungrammatical conditions, a third partition was made: The L2ers performed worse in copula be (C-PSS U) and in thematic verbs (T-SSP U) (around less than 45%) than in copula be (C-SSP) and in thematic verbs (T-PSS U) (around less than 30%). T-PSP G seems to be an exception in the latter category.

Figure 4

Classification Tree Predicting Error Rates in S-VA Agreement with Copulas and Thematic Verbs for both Groups



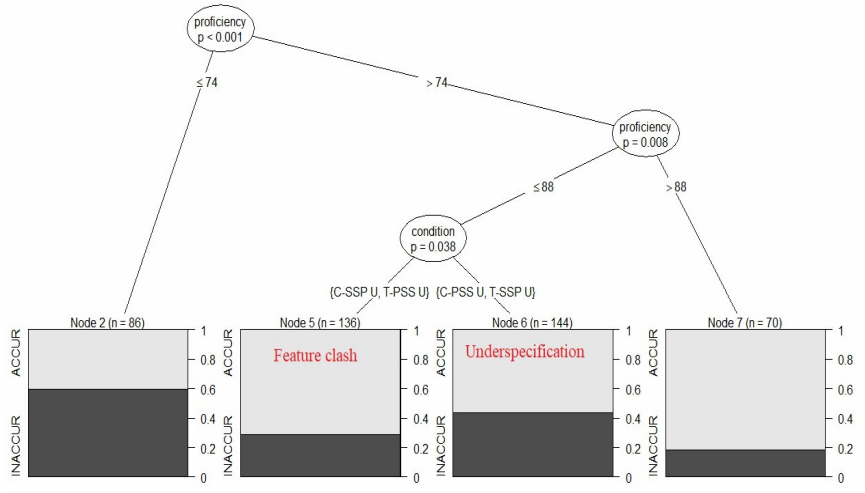
		Thematic verb		Copula *be*
GRAMM	T-SSS G	The boy in the house quarrels frequently.	C-SSS G	The slogan on the poster is expressive.
	T-PSP G	The boys in the house quarrel frequently.	C-PSP G	The slogans on the poster are expressive.
UNGRAMM	T-SSP U	*The boy in the house quarrel frequently.	C-SSP U	*The slogan on the poster are expressive.
	T-PSS U	*The boys in the house quarrels frequently.	C-PSS U	*The slogans on the poster is expressive.

The effect of proficiency is further investigated by running a binary recursive partitioning (the classification tree in Figure 5) to see how proficiency affects the *ungrammatical* conditions (agreement violations). Note that the binary recursive partitioning (Figure 4) showed that L2ers did not show difficulty in *grammatical* conditions.

As Figure 5 shows, the L2ers who scored below 74 (= intermediate L2ers) did not make a difference between conditions; they were generally worse in all conditions than those who scored above 74 (above-74 group). The above-74 group was further divided into two subgroups. Those who scored within the 74-88 range (= advanced L2ers) showed differential sensitivity to ungrammatical conditions, but those scored above 88 (= very advanced L2ers) did not show difficulty in these conditions (only around 10% errors). The 74-88 (advanced) group made worse judgment decisions in (C-PSS U) and (T-SSP U) (around less than 45%) than in C-SSP U and T-PSS U (around less than 30%).

Figure 5

Classification Tree Predicting Error Rates in S-V Agreement with Copulas and Thematic Verbs for L2ers, Proficiency Effect



		Thematic verb		Copula *be*
UNGRAMM	T-SSP U	*The boy in the house quarrel frequently.	C-SSP U	*The slogan on the poster are expressive.
	T-PSS U	*The boys in the house quarrels frequently.	C-PSS U	*The slogans on the poster is expressive.

Summary

With two questions, this study explored L2ers' morphological sources of difficulty in S-V agreement with different verb types (copulas vs. thematic verbs) when participants were under processing pressure: 1. Are L2ers more sensitive to errors of feature clash than underspecification? 2. Does proficiency modulate their sensitivity to these error types? The results revealed that the NSs did not have any problems with S-V agreement in all configurations of both verb types. As for L2ers, there was a significant difference between L2 proficiency subgroups (intermediate,

advanced & very advanced) as exhibited in Table 6.

a. The *intermediate* L2ers were not sensitive to errors of underspecification nor errors of feature clash.

c. The *advanced* L2ers were sensitive to errors of feature clash but not to errors of underspecification.

d. The *very advanced* L2ers converged on the NSs in exhibiting a native-like performance; they were sensitive to both error types.

Table 6
Effect of Proficiency on L2ers' Sensitivity to Errors.

		Intermediate	Advanced	Very advanced
Sensitivity to error type	Feature clash	-	+	+
	Underspecification	-	-	+
Decreased sensitivity = - ; Increased sensitivity = +				

Discussion

Due to the effect of proficiency on the L2ers' sensitivity to agreement errors, the detailed effect of proficiency is discussed in two sections: *intermediate vs. very advanced* and *advanced*.

Intermediate vs. Very Advanced L2ers

The above proficiency levels, shown in Table 6, reveal two opposite ends in L2ers' sensitivity to S-V agreement; whereas the *intermediate* L2ers were not sensitive to either error type (underspecification & feature clash), the *very advanced* L2ers were sensitive to both. The intermediate L2ers' *in*sensitivity to both error types is consistent with L2 processing research (e.g., Ojima et al., 2005; Wakabayashi, 1997) in which intermediate L2ers were not sensitive to errors of feature clash. As for the very advanced L2ers, these observations are consistent with L2 processing research showing native-like performance in highly proficient L2ers (e.g., Armstrong et al., 2018; Rossi et al., 2006; Tanner et al., 2012).

Advanced L2ers

The advanced L2ers demonstrated the omission-commission asymmetry in thematic verbs and the singular-plural asymmetry in copulas. This means that they were more sensitive to errors of feature clash than errors of underspecification in both verb types.

Thematic Verbs: The Omission-Commission Asymmetry. Consistent with the omission-commission asymmetry in thematic verbs, high proficiency L2ers were sensitive to errors of feature clash, as in **Turtles moves slowly*, observed in Ojima et al. (2005) and Armstrong, Bulkes, and Tanner (2018). Likewise, in Wakabayashi's (1997) study, L1 Japanese advanced L2ers of English showed sensitivity to the ungrammatical condition containing errors of feature clash (i.e., overuse condition), as in *The teacher thinks the students like/*likes discussions ….* In contrast, high proficiency L2ers were not sensitive to errors of underspecification in Sato and Felser's (2010) (**She seriously agree*) and Wakabayashi's (1997) studies. This may indicate that advanced proficiency L2ers' response pattern diverged from NSs' when thematic verbs were underspecified default forms.

The above observation is consistent with the prediction of the MUH regarding the omission-commission asymmetry. Likewise, the findings of the

current study support the predictions of the MUH that proficiency is a key factor in enhancing L2ers' ability to be sensitive to errors of feature clash. That is, L2ers at the advanced but not the intermediate levels are sensitive to errors of feature clash.

Copula *be*: The Singular-Plural Asymmetry. Several recent SLA studies showed such a singular-plural asymmetry in intermediate-to-advanced L2ers of L2 Spanish (McCarthy, 2007; 2008; 2012). Additionally, for L2 English several studies (Jiang, 2004; Lim & Christianson, 2015; Tanner et al., 2012) showed that L2ers did not have difficulty when critical items involved *plural marked are*. Thus, L2ers in the current study displayed the same singular-plural asymmetry, observed in the previous studies, despite the presence of intervening elements between subjects and verbs. The singular-plural asymmetry in L2ers' data is consistent with the MUH following the notion of unmarkedness in terms of the Feature Geometric Hypothesis.

In sum, while the intermediate L2ers were not sensitive to either error type, the very advanced L2ers performed at the near-native level. As for the advanced L2ers, they were more sensitive to errors of feature clash than to those of underspecification. As a conclusion, it may be maintained that even though L2ers are likely to have problems with errors of underspecification more than errors of feature clash, proficiency seems to be a key factor modulating L2ers' errors with S-V agreement.

A Representational Deficit in Syntax or Morphology?

The presence of the omission-commission and singular-plural asymmetries in thematic verbs and copulas respectively in L2ers' data may indicate a representational deficit in morphology but not syntax due to the following reasons. First, had the L2 variability been due to the different syntactic raising possibilities between these verb types (cf. Figure 1), it should have led to a suppletive-affixal asymmetry, but it did not. Second, since it is a representational difference between morphological features in geometric hypothesis's terms (Harley & Ritter, 2002) which accounts for the discrepancy in error rates for both verb types, this shows a representational deficit in morphology but not syntax. Specifically, since the underspecified forms have simpler morphological representations than specified ones, they are used as defaults irrespective of their syntactic raising possibilities. Third, as McCarthy (2008) correctly argued, SLA research shows that morphological deficits may not necessarily depend on syntactic ones. For instance, Lardiere (2007) showed that Patty, an L1 Chinese L2ers of L2 English, produces morphological errors (3rd person singular present tense –s and regular past tense –ed) after many years of immersion even though she has long since successfully acquired the corresponding syntactic features. Fourth, since in the reading-based task in this study which requires a bottom-up encoding, lexical forms are temporarily processed sooner than their corresponding abstract syntactic features, morphological effects must have been more influential than syntactic effects in comprehension compared to production tasks, hence this can be further evidence for a morphological deficit but not a syntactic one. Finally, even frequency may not account for the singular-plural asymmetry observed in S-V agreement with copulas because L2ers performed better in *singular subject NP + plural copula be configurations (errors of feature clash) than in *plural subject NP+ singular copula be ones (errors of underspecification). This performance profile is against the corpus data which demonstrate that "the token frequency of 'is' is much larger than that of 'are' in NSs' spoken English" (Safaie, 2015, p. 85). For instance, according to the

COCA database (Davies, 2008), the Subject NP $_{SG}$ + is configuration (e.g., the student is) demonstrates a co-occurrence frequency of 12.23 (per million) words in English. By contrast, the Subject NP $_{PL}$ + are configuration (e.g., the students are) shows a co-occurrence frequency of 0.18 (per million) words in the corpus. This shows that configurations with the singular copula be (i.e., Subject NP $_{SG}$ + is) are even more frequent in English than its plural form (i.e., Subject NP $_{PL}$ + are). Accordingly, had frequency impacted L2ers' performance profiles, L2ers should have exhibited stronger sensitivity to is (the higher frequency forms) than are (the lower frequency ones), but they did not. Similarly, Safaie (2020, 2021) also found anti-frequency effects in high proficiency L2ers' reactions to ungrammatical English regular past tense forms of high-frequency relative to low-frequency ones (i.e., L2ers performed better in detecting ungrammatical forms of low-frequency regular verbs than high-frequency ones).

Proficiency Effect

Proficiency led to the formation of three subgroups: *intermediate*, *advanced,* and *very advanced* L2ers who performed differently. Assuming that "L2 speakers' representations are, in some cases and particularly at lower proficiency levels, deficient" (McCarthy, 2008, p. 483) we may expect variability across proficiency levels. Here, I agree with McCarthy (2008) that the effect of different proficiency levels may show that some L2ers, presumably the intermediate ones in the current study may lack the dependent feature [plural] in their geometries, whereas others, presumably the higher-proficiency ones, may have it. Thus, before acquiring [plural], *intermediate* L2ers' geometries may not instantiate feature asymmetry. Accordingly, at this level, errors are bidirectional as neither singular nor plural is specified, hence errors of underspecification are expected in both directions. In contrast, since the *advanced* ones have already acquired the marked feature [plural], a feature asymmetry causes unidirectional errors (McCarthy, 2008). However, the *very advanced* group did not show difficulty in detecting either error type and behaved like NSs. That is because these L2ers have already acquired the feature asymmetries. But they must also have gained high control over using feature asymmetries as like as NSs, where needed. More specifically, since the GAJ task involved detection of the agreement violations the *very advanced* L2ers must have developed enough control over their feature asymmetries such that they do not allow errors of underspecification nor those of feature clash to occur.

Conclusion and Pedagogical Implications

This study demonstrated morphological properties of L2 errors in S-V agreement. These properties support the MUH which relates L2 variability to a representational deficit in morphology claiming that L2 errors involve systematic substitutions of unmarked forms as defaults in both comprehension and production. That is because L2 errors were driven by systematic variations in morphological markedness under the Feature Geometric Hypothesis (Harley & Ritter, 2002). This theory considers nonfinite thematic verbs and the third person singular form of copula *be* as unmarked or default (underspecified), hence L2ers' reduced sensitivity to errors of underspecification. However, finite thematic verbs and the third person plural form of copula *be* are considered as marked (specified), hence L2ers' increased sensitivity to errors of feature clash. Consequently, L2 variability is a deficit in morphological representations (but not syntactic ones) where unmarked

forms substitute marked forms. Moreover, proficiency plays a key role in developing morphological competence such that, compared to lower proficiency L2ers, higher proficiency ones become sensitive to both marked and unmarked forms.

The least implication for language learning is that L2ers may benefit from a morphologically-oriented approach to language learning more than syntactically-oriented approaches focusing on grammatical accuracy. That is because the syntactic aspect of S-V agreement like hierarchical feature checking is, perhaps, a part of universal principles governing human cognition which may not need to be the centre of instruction. More importantly, the different effects of proficiency levels on L2ers' sensitivity to S-V agreement errors may demonstrate a developmental pattern of language learning more or less similar to the L1 language acquisition pattern. This is because the results of this study along with the existing research were comparable with the developmental pattern of feature geometries in L1 acquisition.

References

Allan, D. (1992). *The Oxford Placement Test* (2nd Edition). Oxford University Press.

Armstrong, A., Bulkes, N., & Tanner, D. (2018). Quantificational cues modulate the processing of English subject-verb agreement by native Chinese speakers: An ERP study. *Studies in Second Language Acquisition, 40*(4), 731-754. https://doi.org/10.1017/S0272263118000013

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Blom, E., & Baayen, H. R. (2012). The impact of verb form, sentence position, home language and proficiency on subject-verb agreement in child L2 Dutch. *Applied Psycholinguistics, 34*(4), 777-811. https://doi.org/10.1017/S0142716412000021

Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism: Language and Cognition, 10*(2), 161-174. https://doi.org/10.1017/S136672890700291X

Cowper, E. (2005). The geometry of interpretable features: Infl in English and Spanish. *Language, 81*(1), 10-46. http://www.jstor.org/stable/4489853

Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990-present. Retrieved from Available online at http://corpus.byu.edu/coca/

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psycholinguistic study. *Studies in Second Language Acquisition*, 27, 141-172. https://doi.org/10.1017/S0272263105050096

Eubank, L., & Grace, S. T. (1998). V-to-I and inflection in non-native grammars. In M.-L. Beck (Ed.), *Morphology and its interfaces in L2 knowledge* (pp. 69-88). John Benjamins.

Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116-124. https://doi.org/10.3758/BF03195503

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender.* John Benjamins.

Franck, J., Lassi, G., Frauenfelder, U., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition*, 101, 173-216. https://doi.org/10.1016/j.cognition.2005.10.003

Harley, H. (1994). Hug a tree: deriving the morphosyntactic feature geometry. In A. Carnie, H. Harley, & T. Bures (Eds.), *MIT Working Papers in Linguistics* (Vol. 21, pp. 289–320).

Harley, H., & Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language, 78*(3), 482-526. https://doi.org/10.1353/lan.2002.0158

Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional feature hypothesis'. *Second Language Research, 13*(3), 187-226. https://doi.org/10.1191/026765897671476153

Herschensohn, J. (2001). Missing inflection in L2 French: Accidental infinitives and other verbal deficits. *Second Language Research, 17*(3), 273-305. https://doi.org/10.1177/026765830101700303

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15(3), 651–674.

https://doi.org/10.1198/106186006X133933

Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'?: Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research, 18*(2), 95-136. https://doi.org/10.1191/0267658302sr195oa

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*(4), 603-634. https://doi.org/10.1017/S0142716404001298

Lago, S., & Felser, C. (2018). Agreement attraction in native and non-native speakers of German. *Applied Psycholinguistics*, 39(1), 619–647. https://doi.org/10.1017/S0142716417000601

Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Mahwah, NJ: Erlbaum.

Lim, J. H., & Christianson, K. (2015). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics, 36*(6), 1283-1315. https://doi.org/10.1017/S0142716414000290

López Prego, B., & Gabriele, A. (2014). Examining the impact of task demands on morphological variability in native and non-native Spanish. *Linguistic Approaches to Bilingualism*, 4(2), 192–221. https://doi.org/10.1075/lab.4.2.03lop

McCarthy, C. (2007). *Morphological variability in second language Spanish* [Unpublished doctoral dissertation]. McGill University, Montreal. https://escholarship.mcgill.ca/concern/theses/hq37vs79g

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research, 24*(4), 459-486. https://doi.org/10.1177/0267658308095737

McCarthy, C. (2012). Modeling morphological variation and development: Person and number in L2 Spanish. *Linguistic Approaches to Bilingualism, 2*(1), 25-53. https://doi.org/10.1075/lab.2.1.02mcc

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. Journal of *Memory and Language, 55*(3), 381-401. https://doi.org/10.1016/j.jml.2006.06.006

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 134-157. https://doi.org/10.1037/0278-7393.21.1.134

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience, 17*(8), 1212-1228. https://doi.org/10.1162/0898929055002436

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*(3), 427-456. https://doi.org/10.1006/jmla.1999.2653

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16*(2), 103-133. https://doi.org/10.1191/026765800677556046

R Development Core Team. (2017). *R: A language and environment for statistical computing: R Foundation for Statistical Computing*, Vienna, Austria. Retrieved from http://www.R-project.org

Radford, A. (2009). *Analysing English Sentences: A Minimalist Approach*. Cambridge: Cambridge University Press.

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18*(12), 2030-2048. https://doi.org/10.1162/jocn.2006.18.12.2030

Safaie, E. (2015). *Sensitivity to subject-verb agreement in second language sentence processing: Evidence from L1 Persian speakers of L2 English* [Unpublished doctoral dissertation]. The University of Essex, Colchester. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.701425

Safaie, E. (2020). Frequency effects of regular past tense forms in English on native speakers' and second language learners' accuracy rate and reaction time. *Journal of Teaching Language Skills, 39*(1), 153-199. https://doi.org/10.22099/jtls.2020.37229.2828

Safaie, E. (2021). Sensitivity to regular and irregular past tense morphology in native speakers and second language learners of English: Evidence from intermediate-to-advanced Persian speakers of L2 English. *Journal of Psycholinguistic Research, 50*, 1107-1135. https://doi.org/10.1007/s10936-021-09790-3

Sato, M., & Felser, C. (2010). Sensitivity to morphosyntactic violations in English as a second language. *Second Language, 9,* 101-118. https://doi.org/10.11431/secondlanguage.9.0_101

Schütze, C. T. (1996). *The empirical base of linguistics: Grammatical judgments and linguistic methodology*. Chicago: University of Chicago Press.

Slabakova, R. (2009). How is inflectional morphology learned? *EuroSLA Yearbook*, 9(1), pp. 56-75. https://doi.org/10.1075/eurosla.9.05sla

Tanner, D., Nicol, J., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds.), *Proceedings of the 36th Boston University conference on language development* (pp. 594-606). Cascadilla.

Wakabayashi, S. (1997). *The acquisition of functional categories by learners of English.* [Unpublished doctoral dissertation]. The University of Cambridge, Cambridge. https://www.researchgate.net/publication/34505048_The_acquisition_of_functional_categories_by_learners_of_English

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

The Effect of Flipped Learning on English Writing Performance and Self-Efficacy of Iranian Medical Students

Research Article
pp. 161-182

Sepideh Nourinezhad[1]
Ehsan Hadipourfard*[2]
Mohammad Bavali[3]

Received: 2020/12/04 Accepted: 2021/10/26

Abstract
The current study investigates the effect of flipped instruction on writing self-efficacy and writing performance of Medical students at Shiraz University of Medical Sciences. Fifty students participated in this experimental study. They were assigned as the treatment group consisted of 25 students and the control group 25 students. The control group (the non-flipped classroom) was taught using traditional writing instruction, whereas the experimental group (the flipped classroom) was taught in a flipped learning mode. The questionnaire used in this study was adapted from Prickel's research (1994) and was scored based on the Likert scale for the writing self-efficacy. The data were gathered in a Pretest-Treatment-Posttest design. The results revealed that flipped instruction had a more positive effect on improvement of both writing self-efficacy and writing performance of the learners compared to those instructed traditionally. The results in this study extended the view point of EFL teachers to understand novel methods of instruction.

keywords: traditional instruction, flipped instruction, writing performance, self-efficacy, writing self-efficacy

*Corresponding Author
[1] PhD Candidate, Department of English, Shiraz Branch, Islamic Azad University, Shiraz, Iran. nouri1542@yahoo.com
[2] Assistant Professor, Department of English, Shiraz Branch, Islamic Azad University, Shiraz, Iran. e.ehadipour@gmail.com
[3] Department of English, Shiraz Branch, Islamic Azad University, Shiraz, Iran. mbvl57@gmail.com

DOI: 10.22051/LGHOR.2021.34132.1409

Introduction

Writing is an essential skill that facilitates communication among human beings. In learning English, writing has been observed as the most complicated skill (Hengwichitkul, 2009; Rattanadilok Na Phuket & Othman, 2015; Watcharapunyawong & Usaha, 2013). Lindemann (1982) refers to writing as a process of communication in which a writer uses the traditional system to show an idea or a content to a reader; to this end, writing is a process in which the writer can convey the message by using letters, punctuation, words or sentences in a graphical system. The process of communication would not be successful unless the reader and the writer understand the language being used in written form. According to Jahin and Idrees (2012), writing seems complex and hard for both native and non-native learners since writers should make a normal balance between numerous issues like purpose, content, audience, mechanics, organization, and vocabulary in their writing drafts.

The present situation of Iranian EFL learners' writing does not seem to meet their needs. Due to the limited exposure of EFL learners to English in Iran which is confined to only language classes (Kafipour et al., 2018), teachers are most of the time concerned about delivering efficient writing instruction to learners. Although the research in the field of writing has been growing recently, it is still far from being satisfactory. Therefore, any studies dealing with students' learning of writing skills, especially in EFL contexts would be of great benefit in the field of language. Due to the advent of technology, it is believed that applying multimedia technology in English teaching will be expanded worldwide; consequently, the significance of this study is incorporating technology in teaching, in this case providing audiovisual aids which might pave the way for both teachers and students to alleviate the afore-mentioned problems. Nowadays, with the great help of computer technology (Computer Assisted Language Learning), it is hoped that the main obstacles of the language learners are identified, and the corresponding remedial teaching methods are put forward. Considering the rapid explosion of science and technology, the vital role of multimedia technology and its function in teaching, in creating audio, visual, and animation effects has been more highlighted in English teaching classes and sets a good opportunity for examination of new models of English teaching in this new era. Several studies have depicted the positive function of multimedia technology in enhancing the students' motivation, and activities in English class. Technological development, digital innovations, and the growth of English have gone hand in hand that are changing the way we communicate (Hekmatzadeh et al., 2016).

It is worth mentioning that the expansion of the Internet has assisted the development of the English language. Regarding this fact, there has been a very crucial increase in the literature considering the application of technology in teaching the English language (Eshet, 2004, 2007; Koehler & Mishra, 2009). Thus, it is necessary for the language instructors to be aware of the latest equipment and have full knowledge about what is available in different situation. Many techniques can be used in various stages of the language learning situation. Some of the techniques in language learning are applicable in testing and distance education, and some are useful for teaching business English, reading, listening, or even interpreting. There are different reasons for knowing how to make use of the new technology. Here, it is also worth mentioning that the new technologies develop in such an accelerating rate that we cannot disregard their power in any form. In this

regard, to overcome the aforementioned difficulties, new technologies are available to be utilized in modern styles, which suit not only the auditory, but also the visual styles of the students (Halwani, 2017; Khojasteh et al., 2016; Tajallizadeh Khob & Rabi, 2014).

Although the research on the writing process is growing, only a few studies have described the usefulness and effectiveness of different strategies, technologies, and even new methodologies accepted and administered by academic writers, though a lot of research (e.g. Brick & Holmes, 2014; Henderson & Phillips, 2014) have focused on the function of digital feedback in removing these problems. The number of studies such as that conducted by Khojasteh et al. (2018) that have explored methods to improve writing instruction through the use of technology with the aim of facilitating teacher-created feedback in English language classrooms is limited (Henderson & Phillips, 2015).

There is a large bulk of literature about evaluation, assessment, and feedback (e.g. Carless & Boud, 2018; Evans, 2016; Evans, 2013; van Heerden et al., 2017), but there is only a small portion, which concentrates on the innovation of digital or computer-based audio-visual instruction, a view shared by Henderson and Phillips (2015) with a lesser amount of literature focused on student achievement. This study will assess the effectiveness of this mode of instruction in terms of student writing performance. Thus, different methods have been used to understand whether other modes of instruction such as audio and video can have any effects on the learners' performance, particularly writing performance. One of the causes of this progress might be increased attention toward distance learning in modern life.

Writing Self-Efficacy and Writing Performance

Self-efficacy refers to trusting our capabilities to face challenging tasks (Gaumer Erickson & Noonan, 2016). Pajares (1996) indicated that self-efficacy influences different factors such as academic motivation, learning, and achievement. Bandura (1977) claimed that people with low self-efficacy easily give up while they encounter difficulties. Thus, estimating the learners' level of academic self-efficacy during their learning procedure is essential (Bandura, 1986b; Schunk, 1985; Zimmerman & Martinez-Pons, 1990). Based on the National Institute of Education (1980), it is necessary to estimate such self-efficacy to improve both instruction and assessment in adult learners' learning procedures.

Beach (1989), Meier et al. (1984), McCarthy et al. (1985), Shell et al. (1989), Zimmerman & Martinez-Pons (1990) have all examined the impact that self-efficacy may have on learners' function in writing, and all of them have found out that self-efficacy is a factor that can predict the actual function of participants in writing. In the same line, Pajares & Valiante's (1996) results revealed that participants' writing self-efficacy in elementary schools depicted their writing performance.

Flipped Instruction

As Bartlett (1994) stated, the traditional method of writing instruction is revealed as a teacher-controlled approach. It refers to traditionally long-established pedagogical procedures, which focus on materials such as pre-printed textbooks (Funk & Funk, 1989). Within the traditional classroom, pupils do not have access to the subject directly; they have little information about the material and issues in the classroom, and they mostly have to do homework after the class at home. In the

classroom, instructors teach all the knowledge, which is often the basic knowledge about an issue, and students have to memorize and learn them at home, which is homework.

Before graduation, learners must attain proficient writing skills. However, based on Bless (2017), the performance of some learners is not satisfactory and acceptable in writing; as a result, academic instructors recommended that a significant aspect to improve the learners' writing skills is to create new methods to apply in classroom rather than the traditional one. This procedure can help the students to promote their overall writing performance and assist the instructors to change and improve their writing instruction. Consequently, many investigations have been conducted on the application and efficacy of a new approach to teaching named the flipped classroom pedagogy (Soliman, 2016; Szparagowski, 2014).

Flipped instruction is providing instructions in a digital format; it is usually a video to be seen before class, and instead, use the class time for communicative group practices. In English language learning and teaching, several studies regarded the vital role of having flipped classrooms. The advantages of applying the flipped model of instruction in language learning classes are increasing autonomous learners (AlJaser, 2017), decreasing the number of formal class meetings (Johnson, 2013), raising learners' ingenuity (Al-Zahrani, 2015; Song & Kapur, 2017), expanding their class activities (Basal, 2015), boosting the participants' stimulus (Davies et al., 2013; Elian & Hamaidi, 2018; Sirakaya & Ozdemir, 2018; Villanueva, 2016), improving the learners' academic performance (Evseeva & Solozhenko, 2015), paving the way for class discussion (Marlowe, 2012), and enhancing the learners' communicative skills (Unakorn & Klongkratoke, 2015). Therefore, the Flipped Classroom is a new pedagogical method that transfers lecturing and note-taking outside the classroom to be done by learners, which enables constant contact between teachers and learners and even learners themselves.

Based on Allen et al. (2011), "Visual aids add interest to an instruction, and individuals can make use of different senses" (p 1-5). In the same vein, as witnessed by Mathew and Alidmat (2013), audio-visual aids would facilitate comprehension.

According to Sugar et al. (2010), audiovisual aids allow the teachers to form behaviors and actions and allow the learners to view and review the content any time and every time they need at their convenience.

What appears to be at the heart of this mode of instruction is the fact that nowadays the students have completely different expectations when compared to the past. Moreover, it is too difficult to draw their attention through the traditional teaching procedures. Furthermore, it is hard to deal with some difficulties regarding teaching and learning via traditional methods. Accordingly, the instructors today must try to create new teaching procedures that meet the needs of this generation; therefore, alternatives to traditional instructions and approaches would be critical. Although it facilitates video recording along with narration comments, the empirical studies investigating the use of audio and visual equipment to provide feedback in EFL contexts are rather limited.

Despite the teachers' attempts to help the learners improve their writing skills, students usually hesitate to write due to the perceived difficulty of foreign language writing. To overcome this problem, teachers have suggested a variety of techniques. To activate the students' digital competencies, teachers can use flipped instruction which is the most noticeable way of shifting from traditional pen-and-paper-based writing classrooms into a more innovative model. Therefore, it is

obvious that traditional writing instruction needs to be changed. This study suggests a new model in teaching writing that is flipped instruction in writing classes.

Literature Review

Flipped Instruction, Having self-Efficacy in Writing, and Writing Performance

In a research conducted by Elian & Hamaidi (2018), the effect of the flipped instruction on fourth-grade science students' academic achievement was explored. The findings revealed that the pupils' function in the experimental group was better in the educational achievement test.

In another research, Sirakaya & Ozdemir (2018) investigated the influence of a flipped mode of instruction on the participants' academic attainment, their self-guided learning, and motivational factors. They came up with the point that the flipped instructed group outperformed the control group considering academic attainment, motivation, and retrieval.

As witnessed by Zheng et al. (2018), using an integrated pedagogical tool, the combination of off-line and flipped classroom activities, showed superior learning results, professional knowledge, and promoted the students' capabilities.

In Hung's (2015) study, the participants who were the receivers of flipped lessons showed a statistically considerable development in their academic functions, which might due to the fact that the teacher-created pre-class audio-visual materials paved the way for the learners to prepare for in-class practices.

This is in line with the findings of Daniel (2013), who talked about the benefits of applying audio-visual aids in teaching English. Some of the advantages of applying this mode of instruction are that it provides interest and motivation for learning, it saves time and explains the ideas easily and precisely, instructor's burden is decreased, different experiences emerge for the students, it makes learning English easy, and it enhances the pupil's attention in the lesson. With teacher-made video instruction, learners can handle the learning process by themselves (Brick & Holmes, 2008), and become more involved in the papers editing process (Thompson & Lee, 2012).

As O'Malley (2011) declares, using multimedia modes in teaching makes the learners feel relaxed because this mode of instruction is given in a personal manner. Accordingly, Ruffini (2012) realized that pupils preferred to receive audiovisual aids due to their greater flexibility because through digital instruction students could handle their writing process via recording the whole instruction as guidance. This is in line with Stannard's investigation (2008) who indicates that audiovisual aids are helpful for different learning styles and preferences. To have an effective teaching, a meaningful relationship between teachers and students is essential (Campbell et al., 2017).

AlJaser (2017), measured the usefulness of flipped instruction in academic attainment and female college students' self-efficacy, she realized that applying a flipped classroom makes learning more productive, and accordingly, teaching, and lecturing more interesting. Furthermore, Halwani (2017) found out that reading and writing improved when instructors used audio-visual aids, and multimedia helped the learners to grasp the issue and became interactive in the classroom with no fear of having problems due to shyness. However, there have been some reports that revealed no significant differences in terms of academic outcome between flipped and non-flipped instruction (Adnan, 2017; Guidry et al., 2013; Kissau et al., 2010), but the researches demonstrated that flipped classrooms should be considered equally

in pedagogical decisions since they have been found to be as effective as traditional modes of instruction.

On the other hand, as claimed by Bandura (1986a), one's behavior, under the influence of self-efficacy, can influence his or her academic success. Based on several studies, there is a relevance between the self-efficacy of participants and their functions in writing (Amogne, 2008; Chen & Lin, 2009; Erkan & Saban, 2011, Shah et al., 2011; Woodrow, 2011); consequently, self-efficacy has a significant influence in the students' writing skills. Several studies were in line with Bandura's assertion and depicted the influence of self-efficacy on learning (e.g. Lane & Lane, 2001; Lane et al., 2004; Pajares & Johnson, 1996; Pajares & Valiante, 1999; Shell et al., 1989; Shell et al., 1995). Based on Chen's (2007) investigation, self-efficacy predicts the students' language performance.

Daly in 1978 examined 3602 undergraduate students in writing classes with different levels of writing apprehension. He found out that those with little writing anxiety had better writing skills compared to learners with high apprehension. The results of this study are congruent with Lee and Krashen (1997), who conducted a study in Taiwan on the native speakers of Chinese. The findings revealed that students who have higher writing anxiety did not have any tendency to receive assessment. In another study with a similar type of outcomes, Woodrow (2011) explored the significance of self-efficacy in 738 Chinese university participants' writing. Like several other studies, the results depicted that an important link was between writing self-efficacy and writing performance.

In another research, Shah et al. (2011) worked on 120 Malaysian participants' general self-efficacy and their functions in writing. He concluded their test score in writing had a high positive relationship with their self-efficacy in writing.

As mentioned before, the present situation of Iranian EFL learners' writing does not seem to meet the needs; although a considerable amount of time is assigned in developing EFL writing, the result is still disappointing. Therefore, any research dealing with students' writing skills, especially in EFL contexts, would be of great benefit in the field of language. In the area of foreign language teaching, based on Lee (2003), a key factor for instructors, researchers, textbook writers, and program designers is having a good EFL writing. In the EFL context in Iran where the current study is conducted, research has shown that the flipped instruction, as an innovative mode in teaching procedure, depicted reassuring outcomes for language learning. For example, Mohammadi et al. (2019) came up with the point that flipped classrooms caused the Iranian EFL learners' language competence to enhance. Moreover, flipped instruction has been confirmed to have effective outcomes in Iranian EFL learners' writing skills (Abedi et al., 2019). Consequently, this research was conducted to better understand the impact of flipped instruction model which may influence the Iranian writing self-efficacy and writing performance, considering the point that both Iranian EFL teachers and learners have the tendency towards applying flipped mode of instruction in language teaching and learning (Jafarigohar et al., 2019; Vaezi et al., 2018).

However, there were some other research projects with different outcomes. For example, Al-Mekhlafi's investigation (2011) on the self-efficacy of Arab EFL trainee-teachers' in writing and their writing attainment contradicted the findings of the aforementioned studies. It depicted that there was no relationship between the

two variables. In this study, the participants' overall writing score includes portfolio, mid-semester test, class activity, and final scores affected self-efficacy since it is a task-specific variable. According to the literature review and considering the development of technology, the need to investigate the effect of flipped instruction, as a new model of instruction, on writing self-efficacy and writing performance seems urgent, especially in EFL context. The present study sought to answer the following research questions:

1. Does the type of instruction (i.e. flipped vs. traditional) influence the EFL students' writing self-efficacy?
2. Does the type of instruction (i.e. flipped vs. traditional) influence the EFL students' writing performance?

Methodology

Design of the Study

The present research applied an experimental research design to examine two experimental and control groups. We aimed to explore whether using flipped instruction can influence the writing self-efficacy and writing performance of medical students studying at Shiraz University of Medical Sciences.

Participants

The population of this study consisted of all medical students studying at Shiraz University of Medical Sciences in the 2019-2020 fall semester and those who had taken a 3-unit compulsory writing course. Each group consisted of 25 participants (a total of 50 students) who were assigned between two writing classes and their ages ranged from 20 to 24 years. Concerning the experimental nature of this study, researchers did their best to be very careful about any intervention between the results of the study and the environmental aspects that could impact on the study.

Instruments

Questionnaire. The researchers applied the questionnaire from Donald O. Prickel's research in 1994 to investigate the participants' self-efficacy in writing. This questionnaire consisted of 25 Likert scale questions, with 5 options as Strongly Disagree, Disagree, Unsure, Agree, and Strongly Agree. The back-translation method was applied to translate this questionnaire from English into the Persian language. In this regard, a competent translator first translated the questionnaire into Persian. After that, without reference to the original text, another professional translator again converted the Persian form of the questionnaire into English and then the researchers compared these two English versions.

It is good to mention that the researchers tested the content validity of this questionnaire by consultation with 3 experts to know whether each question reflects its intended concepts. Besides, the researchers tested the reliability of the questionnaire by means of Cronbach alpha and .79 was found as the internal consistency coefficient of the questionnaire, which was within the acceptable range. They collected the data in a pretest and posttest treatment design to examine the impact of flipped instruction on participants' self-efficacy in writing as the first dependent variable.

Test Scores. The researcher also used the students' writing scores and explored the effect of flipped instruction on the participants' functions in writing as

the second dependent variable in both classes. To make sure that these students were homogenous in writing ability, we collected and analyzed their very first writing assignment- before treatment- as a pre-test using an independent sample t-test. Based on the results, the difference was not significant; therefore, the two writing groups were homogenous. After 12 weeks, the students' midterm exam data were collected and analyzed as post-test.

Writing Grading Rubric. The Analytic Rubric by Jacobs et al. (1981) was used to score the learners' writing assignments. The rubric assessed the medical students' skill in writing using 5 traits, including content, organization, language use, vocabulary, and mechanics. The scores allocated to each of these traits included: Content = 25, Organization = 25, Language use = 25, Vocabulary = 15, and Mechanics = 10. The total mark was 100 points. The writing rubric can be seen in Appendix A.

Reliability Test. Since the researchers were dealing with the human rater, it was essential and suggested by Neuendorf (2002) to ask at least another rater to rate the written assignment of medical students in two phases of the data gathering. According to Carmines and Zeller (1979), reliability has been defined as a measuring procedure that on repeated actions it produces the same results. Therefore, among all the university lecturers teaching English to medical students at Shiraz University of Medical Sciences, the researcher asked two of them to mark the papers for content, organization, word choice, sentence structure, grammar, and mechanics, as mentioned in the rubric. To make sure that the raters were in the same line with one another, the researcher had a two-hour session practice with the raters to clarify all the elements of the writing rubric. Then, 50 pre-test and 50 post-test assignments were given to the raters and they were asked to return the papers within one month. To prevent subjectivity in marking the assignments, neither the researchers nor the writing instructor of the two groups marked the papers.

Based on the scores of five components in each writing task, the researchers applied Pearson product-moment correlation to quantify the inter-rater reliability. For the components content, organization, word choice, sentence structure, grammar, and mechanics, the reliability of the total score was .678 followed by .681, .647, .743, .618, and .698, respectively. This reveals a high congruity between the two raters considering the process in which both scored the students' performance in writing and its components. Furthermore, the correlation coefficients were significant, demonstrating the consistency of both raters in scoring the participants' writing assignments.

Writing Video Contents Used in the Flipped Instruction. In this study, the researcher used professionally pre-recorded grammar lessons in the Virtual University of Medical Sciences. These lessons were uploaded in the Learning Management System known as LMS and could be accessed only by the students in flipped method group. To make sure that the videos cannot be reached to the students instructed by the traditional method, all videos were un-downloadable, so the students had to access their LMS any time they wanted to watch the videos. It is worth mentioning that the LMS website requires the students' ID number and password.

Procedure

Treatment Group. To explore the possible relationship between flipped instruction and students' writing performance, in the treatment group, the

researchers used teacher-made videos for each session in the course of the whole semester. In this regard, instead of describing everything, the writing instructor used flipped instruction model of teaching. For this study, the instructor asked the students to watch the assigned videos before attending a class. When attending the flipped classroom sessions, the students were required to do class exercises (such as editing paragraphs, rewriting paragraphs, etc.) provided by their writing instructor individually or in groups. In this way, the students themselves covered the writing contents at home, whereas they did the assignments in the classroom unlike what happens in traditional classes.

Control Group. Students in this group received traditional instruction. This model of instruction consisted of teaching all the contents covered by the flipped group; however, these contents were taught to students in the classroom and they were asked to do the assignments at home. In the following sessions, the writing instructor delivered the answers to the participants in the classroom. The control group have a traditional writing class. That is, the participants also went to the class and worked on the similar issues like the experimental group. The only difference was that for the control group, the traditional model was utilized in writing instruction.

Data Analysis

In this study, the researcher applied SPSS, version 20. To deal with both aims, "to investigate whether flipped instruction has any effect on writing self-efficacy" and "to investigate if it influences the students' writing performance", an independent sample t-test was applied, to show if the mean score of the participants in one situation significantly varies from that of the other situation.

Results

Table one shows descriptive statistics including mean and standard deviation for pre-test and post-test scores obtained for writing performance and self-efficacy.

Table 1
Descriptive Statistics

Instruction		N	Mean	Std. Deviation	Std. Error Mean
Pretest for writing performance	traditional	25	48.31	6.57	1.40
	flipped	25	45.64	7.92	1.58
Posttest for writing performance	traditional	25	62.50	9.20	1.96
	flipped	25	95.28	3.84	.76
Pretest for self-efficacy	traditional	25	2.72	.70	.14
	flipped	25	2.64	.63	.12
Posttest for self-efficacy	traditional	22	2.86	.63	.13
	flipped	25	4.20	.47	.09

As shown in Table 1, pre-test mean scores of writing self-efficacy in the traditionally instructed group (m = 2.72) are slightly higher than that of the flipped-instructed group (m = 2.64). To ensure if this difference is statistically significant, the investigators conducted an independent sample t-test. As depicted in Table 2, the Sig. value for writing self-efficacy between pre-test of both groups is greater than .05 (Sig. = .66); therefore, the researchers deduced that no significant difference is

between pre-tests of both traditional and flipped group based on their writing self-efficacy; consequently, the two groups were homogenous considering writing self-efficacy.

To answer the first research question - Does the type of instruction (Flipped versus Traditional) influence the writing self-efficacy? - as shown in Table 1, the post-test mean score of writing self-efficacy for traditional instruction (m = 2.86) was lower than that of flipped instruction (m = 4.2). It shows that flipped-instructed group outperformed the traditionally-instructed one. To ensure if this difference is statistically significant, we conducted independent samples t-test was between the posttest scores of both groups.

Table 2
Independent Samples t-Test

	F	t	df	Sig. (2-tailed)
Pretest writing performance	.002	1.251	45	.217
Posttest writing performance	16.672	-16.282	45	.000
Pretest self-efficacy	.121	.446	45	.657
Posttest self-efficacy	1.202	-8.169	45	.000

As shown in Table 2, Sig. level for post-test writing self-efficacy is .000, which is lower than .05. It reveals a statistically significant difference between post-test mean scores in both groups. Therefore, it can be concluded that flipped instruction improved the participants' writing self-efficacy in comparison with traditional instruction. As shown in Table 1, pre-test mean scores of writing performance in the traditionally-instructed group (m = 48.3) were greater than that of the flipped-instructed group (m = 45.6). The researchers conducted an independent sample t-test to ensure if this difference was statistically significant. As seen in Table 2, the Sig. value for writing performance between the pre-tests of both groups was greater than .05 (Sig. = .21); therefore, it can be inferred from the results that there was no significant difference between the pre-tests of both groups regarding their writing performance, and the two writing groups were homogenous.

To answer the second research question- Does the type of instruction (Flipped versus Traditional) influence the writing performance of selected students? - according to Table 1, the post-test mean score of writing performance for traditional instruction (m = 62.5) was greatly lower than that of flipped instruction (m = 95.2). It shows that participants in the flipped-instructed group performed better than those in the traditionally-instructed group. Yet, to ensure if this difference is statistically significant, we used the independent samples t-test between the post-test scores of both groups.

As shown in Table 2, the Sig. level of the post-test score for writing performance was .000 which is lower than .05. It indicates a statistically significant difference between the post-test mean scores in both groups. Therefore, it can be concluded that flipped instruction has a more positive influence in improving the writing performance in the flipped-instructed group compared to the traditionally-instructed one.

Findings and Discussions

This study focused on the writing skill as one out of 4 language skills due to the problems Iranian students face in second language writing. English writing is a problematic and challenging task for Iranian students (Kafipour et al., 2018); that is why the current study was an attempt to evaluate a possibly more effective approach -flipped instruction in improving the learners' writing skills.

The first research question in this study was: "Does the type of instruction (i.e. flipped vs. traditional) influence the EFL students' writing self-efficacy?" There was an attempt to find out if there is relevance between applying flipped instruction and the medical students' writing self-efficacy. The findings of this study confirmed this relationship. The study demonstrated that flipped instruction in comparison with traditional instruction improved the students' writing self-efficacy. This may be due to the reason that receiving materials before the class and studying it ubiquitously might have permitted the participants to control their own learning to have better learning outcomes. Moreover, as we have seen in this study, flipped instruction provided a new learning model in which learners were the receivers of different self-learning modes. These modes along with a convenient learning environment, which helped them conquer their writing problems can increase their self-efficacy in writing.

This study has revealed similar findings to the study by Hamdam et al. (2013). They asserted, in flipped instruction, "Learners can explore topics in greater depth and create better learning opportunities." (p. 5) Such learning context might create a strong attitude in attaining success, which in return can cause improved self-efficacy.

According to Raimes (1983), pictures and videos are fruitful aids for instructors. Pictures and videos help the students to put appropriate vocabularies and ideas in sentences (Asrifan, 2015; Kurniati, 2015; Styati, 2016). Several scholars have stated that applying pictures and videos enables the students to simultaneously use more than one sense, which are efficient in enhancing their writing skills (Asrifan, 2015; Kaur et al., 2017; Kurniati, 2015; Sesrica & Jismulatif, 2017; Styati, 2016; Wening, 2016). Therefore, variation can be involved in the classroom via different visuals aids, which could enhance the learners' focus on the issues. In the same vein, Mathew and Alidmat (2013) realized that audio-visual facilities make the issues more comprehensible; moreover, they believed that in this way the students could recall the materials for a longer time. Consequently, applying visual aid, in this case flipped mode of instruction, is attractive for promoting the students' achievement.

This study is also consistent with Samiee Zafarghandi (2018). According to the findings of his study, the mean scores of academic achievement and self-efficacy of participants who received the flip learning mode were significantly greater than those who used the traditional model; therefore, the flip learning model was efficient in the learners' self-efficacy and academic achievement.

This is also in line with the study conducted by Lee Su Ping et al. (2020). They found that in this mode of instruction, most students had positive experiences such as greater time preparation before class; increased practice, engagement, interaction, motivation; immediate feedback during class; and a higher level of self-efficacy after class.

The second research question in this study was: "Does the type of instruction (i.e. flipped vs. traditional) influence the EFL students' writing

performance?" The result of this study indicated that flipped instruction significantly improved the students' writing performance. The current results are in accordance with a study conducted in Japan by Leis et al. (2015), who flipped their English writing class to investigate the effectiveness of this mode of instruction. Overall, it has been proven that flipped instruction results in better improvement in the students' writing abilities.

This result is also consistent with the finding of Halwani (2017). He reported that multimedia could help the learners to improve reading and writing skills and interact in the classroom with no fear of having difficulties due to shyness.

This finding supports those obtained by Elian and Hamaidi (2018), who found that flipped instruction significantly improved the students' academic achievement. This result is also supported by another study conducted by Sirakaya and Ozdemir (2018) who came up with the point that flipped instruction enhanced the students' academic achievement, motivation, and retrieving. AlJaser (2017) also found out that female students in her study learned more productively using flipped instruction, while teaching was also interesting for instructors.

The results of the current study were also consistent with those of the study conducted by Zheng et al. (2018). They realized that the integrated pedagogical tool that mixes off-line and flipped classroom activities revealed superior learning outcomes and promoted the students' capabilities.

The findings of this research, considering the participants' writing performance, are also in the same line with the study done by Özkurkudis and Bümen (2019). They understood that the receivers of the flipped instructed group were significantly greater than the control group. They also found that using the flipped model was time-saving and easily accessible.

Based on the results, there were many elements involved in improving students' English writing performance in flipped instruction. The interactive environment, adjustability in both time and place, and diversity in learning sources are among these factors, which are based on Vygotsky's social constructivism (1978). The concept of learning, according to his learning theory, is the students' self-endeavor in creating new knowledge and meaning through social interaction. Flipped instruction provides the students with the essential adjustability to work together, and cooperate without time and place limitations. Consequently, this is highlighted by Kassem (2017); he maintained that "Paring intensive exposure to the learning material outside the classroom with the cooperative in-class activities contributed significantly to the students' high academic achievement." (p. 21)

The results of the present study are also in accordance with other investigations (Alkhoudary, 2019; Qader & Arslan, 2019; Özkurkudis & Bümen, 2019; Tuna, 2017). The students' progress in writing might be due to their knowledge about the writing aspects when they were working together in the flipped classes. In flipped mode of instruction, the students were so motivated for group discussion in which a friendly cooperation was promised (Al-Bahrani, 2020). This interaction can provide students with immediate scaffolding from other learners to get the desired aims in the writing task.

This agrees with the findings presented by Alghasab (2020), who examined the influence of flipping a writing classroom on the writing competencies and perceptions of EFL learners. The results indicated that the EFL students held positive ideas towards the flipped classroom. Further investigation revealed that flipping the writing classroom created a more useful learning environment, provided

flexible instruction, which enhanced the students' writing strategy use, and improved their motivation and interaction.

Based on Bandura's theory in 1977, self-efficacy can influence an individual's academic achievement. This is also similar to those of Shah et al. (2011). They came up with the point that there was a relationship between Malaysian participants' writing skills with their writing self-efficacy. The findings confirmed that this theory held true regarding the application of flipped instruction in writing.

Conclusion

Based on many studies, there is a significant relationship between the learners' self-efficacy beliefs and their writing performance (Amogne, 2008; Chen & Lin, 2009; Erkan & Saban, 2011; Shah et al., 2011; Woodrow, 2011). Accordingly, based on the current study, the researchers concluded that using flipped instruction improves the writing self-efficacy and writing performance of the participants in comparison with traditional instruction. One possible reason is that within the traditional classroom, pupils do not have access to their teacher's explanations and instructions directly at home because learners receive the information only in the classroom and they should do homework at home after the class.

The other thing to consider is that the definition of the term flipped classroom refers to an inversion of the place learning activities occur (Wilson, 2013). Therefore, in this mode of instruction learners have accessed digital materials and their teachers' instruction on different subjects at home. Through this model of instruction, according to Abeysekera and Dawson (2015), "motivation and self-efficacy to learn are improved when environments provide opportunities for students to (a) feel experienced in their capabilities, (b) feel a sense of connection to other pupils and instructors during learning, and (c) be autonomous in self-regulating and decision making." (p. 2)

Furthermore, considering the Bloom's taxonomy, outside the class the learners do the lower levels of cognitive work, and in the class, they focus more on the higher orders (Brame, 2013).

Therefore, based on the aforementioned reasons, concerning the context of the revision process, and the possibility of watching the teacher created instructing videos, in this study it was found that the experimental group who received flipped instruction outperformed in writing self-efficacy and writing performance in comparison with the traditionally-instructed group.

Pedagogical Implications

The results of the present research entail some noteworthy issues in the study of audiovisual instruction. From a pedagogical perspective, this study introduced useful insights for EFL teachers, writing researchers, and learners. The results of this study broaden the instructors' attitudes toward novel methods of teaching and they may use more suitable strategies to improve the students' learning, especially in this era when the world is affected by the Covid-19 pandemic and the importance of remote learning and teaching is highlighted. Accordingly, this mode of instruction is fruitful for EFL students in that they can distinguish which mode is more feasible in the process of learning.

Limitations of the Study

This research could not provide a large sample of participants to show the effect of audiovisual aids and different modes of instruction on the students' writing ability. Future researchers may be able to provide larger samples to offer a wider and more visible image of the effect of AVF plus different modes of instruction on the students' writing ability considering different variables.

References

Abedi, P., Namaziandost, E., & Akbari, S. (2019). The impact of flipped classroom instruction on Iranian upper-intermediate EFL learners' writing skill. *English Literature and Language Review*, *5*(9), 164-172. https://doi.org/10.32861/ellr.59.164.172

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *Higher Education Research & Development*, *34*(1), 1-14. https://doi.org/10.1080/07294360.2014.934336

Adnan, M. (2017). Perceptions of senior-year ELT students for flipped classroom: A materials development course. *Computer Assisted Language Learning, 30*(3-4), 204-222. https://doi.org/10.1080/09588221.2017.1301958

Al-Bahrani, R. H. (2020). Flipped classroom as a new promoting class management technique. *Journal of College of Education/Wasit, 2*(39), 591-610. https://doi.org/10.31185/eduj.Vol2.Iss39.1426

Alghasab, M. B. (2020). Flipping the writing classroom: Focusing on the pedagogical benefits and EFL learners' perceptions. *English Language Teaching*, *13*(4), 28-40. https://doi.org/10.5539/elt.v13n4p28

AlJaser, A. (2017). Effectiveness of using flipped classroom strategy in academic achievement and self-efficacy among education students of Princess Nourah Bint Abdulrahman University. *Canadian Center of Science and Education*, *10*(14), 67-77. https://doi.org/10.5539/elt.v10n4p67

Allen, K., & Marquez, A. (2011). Teaching vocabulary with visual aids. *Journal of Kao Ying Industrial & Commercial Vocational High School, 1*(9), 1-5. http://210.60.110.11/reading/wp-content/uploads/2012/10/10022007.pdf

Alkhoudary, Y. A. (2019). Flipping the classroom towards learner autonomy in an EFL writing class [Paper presentation]. *The 1st applied linguistics and language teaching conference: Teaching and learning in a globalized world* (pp. 47-58). Zayed University Press.

Al-Mekhlafi, M. A. (2011). The relationship between writing self-efficacy beliefs and final examination scores in a writing course among a group of Arab EFL trainee-teachers. *International Journal for Research in Education (IJRE)*, *29*, 16-33. www.fedu.uaeu.ac.ae/journal/docs/pdf/pdf29/2_E.pdf

Al-Zahrani, A. (2015). From passive to active: The impact of the flipped classroom through social learning platforms on higher education students' creative thinking. *British Journal of Educational Technology, 46*(6), 1133-1148. https://doi.org/10.1111/bjet.12353

Amogne, D. (2008). *An investigation of the correlation among efficacy sources, students' self-efficacy and performance in reading and writing skills: Bahir Dar University in Focus* [Unpublished doctoral dissertation]. Addis Ababa University, Addis Ababa, Ethiopia. http://localhost:80/xmlui/handle/123456789/7607

Asrifan, A. (2015). The use of pictures story in improving students' ability to write narrative composition. *International Journal of Language and Linguistics, 3*(4), 244-251. https://doi.org/10.11648/j.ijll.20150304.18

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review, 84*, 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986a). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall. https://www.jstor.org/stable/1449722

Bandura, A. (1986b). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359-373.

https://doi.org/10.1521/jscp.1986.4.3.359
Bartlett, A. (1994). The implications of whole language for classroom management. *Action in Teacher Education, 16*, 65-74. https://doi.org/10.1080/01626620.1994.10463189
Basal, A. (2015). The implementation of a flipped classroom in foreign language teaching. *Turkish Online Journal of Education-TOJCE, 16*(4), 28-37. https://doi.org/10.17718/tojde.72185
Beach, R. (1989). Showing students how to assess: Demonstrating techniques for response in the writing conference. In C. M. Anson (Ed.), *Writing and response*, (pp. 127-148). NCTE.
Bless, M. M. (2017). *Impact of audio feedback technology on writing instruction* [Unpublished doctoral dissertation]. Walden University, Minneapolis, USA. https://opencommons.uconn.edu/nera-2017
Brame, C. (2013). *Flipping the classroom.* Vanderbilt University Center for Teaching. http://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/.
Brick, B., & Holmes, J. (2008, October). *Using screen capture software for student feedback: Towards a methodology using screen capture software for student feedback.* [Paper presentation]. The IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2008), Freiburg, Germany. https://www.researchgate.net/publication/255585177
Brick, B., & Holmes, J. (2014, February). *Using screen capture for student feedback*. [Paper presentation]. The Eleventh International Conference Cognition and Exploratory Leaning in Digital Age, Porto, PT. http://www.iadis.org
Campbell, B. S., & Feldmann, A. (2017). The power of multimodal feedback. *Journal of Curriculum, Teaching, Learning and Leadership in Education, 2*(2), 1-6. https://digitalcommons.unomaha.edu/ctlle
Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354
Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage ublications.
Chen, H. (2007). *The relationship between EFL learners' self-efficacy beliefs and performance* [Unpublished doctoral dissertation]. Florida State University, United States. http://purl.flvc.org/fsu/fd/FSU_migr_etd-3846
Chen, M. C., & Lin H. (2009). Self-efficacy, foreign language anxiety as predictors of academic performance among professional program students in a general English proficiency writing test. *Perceptual and Motor Skills*, *109*(2), 420-430. https://doi.org/10.2466/pms.109.2.420-430
Daly, J. A. (1978). Writing apprehension and writing intensity in business and industry. *Journal of Educational Research*, *72*, 10-14. https://www.jstor.org/stable/27537168
Daniel, J. (2013). Audio-visual aids in teaching of English. *International Journal of Innovative Research in Science, Engineering and Technology, 2(8)*, 3811-3814.
Davies, R., Dean, D., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development - ETR&D*, *61*(4), 563-580. https://www.researchgate.net/publication/256446014
Elian, Sh. & Hamaidi, D. (2018). The effect of using flipped classroom strategy on the academic achievement of fourth grade students in Jordan. *IJET*, *13*(2), 110-125. https://doi.org/10.3991/ijet.v13i02.7816
Erkan, D. Y., & Saban, A. (2011). Writing performance relative to writing apprehension, self-efficacy in writing, and attitudes towards writing: A correlational study in Turkish tertiary-level EFL. *The Asian EFL Journal Quarterly*, *13*(1), 164-192.

Eshet, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, *13*(1), 93-106. https://www.learntechlib.org/primary/p/4793/

Eshet, Y. (2007). Teaching online: Survival skills for the effective teacher. *Inroads- The SIGCSE Bulletin*, *39*(2), 16-20. https://doi.org/10.1145/1272848.1272862

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, *83*(1), 70-120. https://doi.org/10.3102/0034654312474350

Evans, C. (2016). *Enhancing assessment feedback practice in higher education: The EAT framework*. University of Southampton. https://eatframework. org. uk

Evseeva, A., & Solozhenko, A. (2015). Use of flipped classroom teaching in language learning. *Procedia- Social and Behavioral Science*s, *206*, 205-209. https://doi.org/10.1016/j.sbspro.2015.10.006

Funk, G. D., & Funk, H. D. (1989). Roadblocks to implementing the writing process. *Clearing House, 62,* 222-224. https://www.jstor.org/stable/30196919

Gaumer Erickson, A., & Noonan, P. (2016*). College and career competency (CCC) framework needs assessment.* University of Kansas, Center for Research on Learning. http://resources.cccframework.org

Guidry, K. R., Cubillos, J., & Pusecker, K. (2013, May). The connection between self-regulated learning and student success in a hybrid course. In *Association for Institutional Research Annual Forum* (Vol. 31). http://www. mistakengoal. com/docs/Self-regulated_learning_hybrid_course. pdf.

Halwani, N. (2017). Visual aids and multimedia in second language acquisition. *English Language Teaching*, *10*(6), 53-56. https://doi.org/10.5539/elt.v10n6p53

Hamdam, N., McKnight, P., McKnight K., & Arfstrom K. M. (2013). *The flipped learning model: A white paper based on the literature review titled - A review of flipped learning*. https://flippedlearning.org/wpcontent/uploads/2016/07/WhitePaper_FlippedLearning.pdf

Henderson, M., & Phillips, M. (2015). Video-based feedback on student assessment: Scarily personal. *Australasian Journal of Educational Technology*, *31*(1), 51-66. https://doi.org/10.14742/ajet.1878

Hekmatzadeh, M., Khojasteh, L., & Shokrpour, N. (2016). Are emotionally intelligent EFL teachers more satisfied professionally? *International Journal of Applied Linguistics & English Literature*, *5*(2), 1-11. https://doi.org/10.7575/aiac.ijalel.v.5n.2p.97

Hengwichitkul, L. (2009). *An analysis of errors in English abstracts translated by Thai university graduate students* [Unpublished master's thesis]. Srinakharinwirot University, Bangkok, Thailand. https://doi.org/10.5539/elt.v10n3p101

Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning, 28*(1), 81-96. https://doi.org/10.1080/09588221.2014.967701

Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach.* Newbury House Publishers, Inc. https://doi.org/10.1177/026553228400100210

Jafarigohar, M., Khoshsima, H., Haghighi, H., & Vahdany, H. (2019). Incorporation of flipped learning into EFL classrooms performance and perception. *Iranian Journal of English for Academic Purposes*, *8*(3), 1-14.

Jahin, J. H., & Idrees, M. W. (2012). EFL major student teachers' writing proficiency and attitudes towards learning English. *Journal of Taibah University*, *4*(1), 9-72.

Johnson, G. (2013). *Student perceptions of the flipped classroom* [Unpublished MA thesis]. University of British Columbia, Vancouver. https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0073641

Kafipour, R., Mahmoudi, E., & Khojasteh, L. (2018). The effect of task-based language teaching on analytic writing in EFL classrooms. *Cogent Education*, *5*(1), 1-16. https://doi.org/10.1080/2331186X.2018.1496627

Kassem, M. A. M. (2017). Flipping the literature classroom: Fostering EFL students' achievement and autonomy. Wadi El-Nile Journal for Humanitarian, Social and Educational Studies and Research, 14(9), 1-28. https://scholar.google.com/scholar?hl=ar&as_sdt=0%2C5&q=Flipping+the+Literature+Classroom%3A+Fostering+EFL+Students%27+Achievement+and+Autonomy.%E2%80%8F+&btnG=

Kaur, C., Singh, S., Mei, T. P., Abdullah, M. S., Mazlini, W., Nor, O., & Mostafa, A. (2017). ESL learners' perspectives on the use of picture series in teaching guided writing. *International Journal of Academic Research in Progressive Education and Development*, *6*(4), 74-89. https://doi.org/10.6007/IJARPED/v6-i4/3463

Khojasteh, L., Shokrpour, N., & Afrasiabi, M. (2016). The relationship between writing self-efficacy and writing performance of Iranian EFL students. *International Journal of Applied Linguistics & English Literature*, *5*(4), 29-37. http://dx.doi.org/10.7575/aiac.ijalel.v.5n.4p.29

Kissau, S., McCullough, H., & Pyke, J. G. (2010). Leveling the playing field: The effects of online second language instruction on student willingness to communicate in French. *Calico Journal*, *27*(2), 277-297. https://doi.org/10.1558/cj.27.2.277-297

Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education*, *9*(1), 60-70. https://www.learntechlib.org/primary/p/29544/

Kurniati, R. (2015). *The effectiveness of using pictures on students' writing of recount Text (A quasi-experimental study at the Tenth Grade Students of SMA Negeri 9 Kota Tangerang Selatan).* The Department of Education, Faculty of Tarbiyah and Teachers Training, Syarif Hidayatullah Atate Islamic University Jakarta. http://repository.uinjkt.ac.id/dspace/bitstream/123456789/30171/3/ROSYAKURNIATI-FITK.pdf

Lane, J., & Lane, A. (2001). Self-efficacy and academic performance. *Social Behavior and Personality, 29*(7), 687-694. https://doi.org/10.2224/sbp.2001.29.7.687

Lane, J., Lane, A., & Kyprianou, A. (2004). Self-efficacy, self-esteem and their impact on academic performance. *Social Behavior and Personality: An International Journal*, *32*(3), 247-256. https://doi.org/10.2224/sbp.2004.32.3.247

Lee, S. (2003). Teaching EFL writing in the university: Related issues, insights, and implications. *Journal of National Taipei Teachers College*, *16*(1), 111-136.

Lee, S. Y., & Krashen, S. (1997). Writing apprehension in Chinese as a first language. *ITL: Review of Applied Linguistics, 115*(1), 27-37. https://doi.org/10.1075/itl.115116.02lee

Leis, A., Cooke, S., & Tohei, A. (2015). The effects of flipped classrooms on English composition writing in an EFL environment. *International Journal of Computer-Assisted Language Learning and Teaching*, *5*(4), 37-51. https://doi.org/10.4018/IJCALLT.2015100103

Lindemann, E. (1982). *A rhetoric for writing teachers*. Oxford University Press.

Marlowe, C. (2012). *The effect of the flipped classroom on student achievement and stress.* [Unpublished doctoral dissertation]. Montana State University, Bozeman, MN

https://scholarworks.montana.edu/xmlui/handle/1/1790

Mathew, N. G., & Alidmat, A. O. H. (2013). A study on the usefulness of audio-visual aids in EFL classroom: Implications for effective instruction. *International Journal of Higher Education*, *2*(2), 86-91. https://doi.org/10.5430/ijhe.v2n2p86

McCarthy, P., Meier, S., & Rinderer, R. (1985). Self-efficacy and writing. *College Composition and Communication*, *36*(4), 465-471. https://doi.org/10.2307/357865

Meier, S., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research*, *8*, 107-120. https://doi.org/10.1007/BF01173038

Mohammadi, J., Barati, H., & Youhanaee, M. (2019). The effectiveness of using flipped classroom model on Iranian EFL learners' English achievements and their willingness to communicate. *English Language Teaching*, *12*(5), 101-115. https://doi.org/10.5539/eltv12n5p101

National Institute of Education. (1980). *National council on educational research. Fifth report: Fiscal years 1978-1979*. National Council on Educational Research No. ED 215-445.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Sage. http://lccn.loc.gov/2015044657

O'Malley, P. J. (2011). Combining screen casting and a tablet PC to deliver personalized student feedback. *New Directions, 7*, 27-30. https://doi.org/10.29311/ndtps.v0i7.464

Özkurkudis, M. J., & Bümen, N. T (2019). Flipping the writing classroom: Using grammar videos to enhance writing. *Journal of Education and Future*, *15*, 1-16 https://doi.org/10.30786/jef.425632

Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research*, *66*, 543-578. https://doi.org/10.3102/00346543066004543

Pajares, F., & Johnson, M. J. (1996). Self-efficacy beliefs and the writing performance of entering high school students. *Psychology in the Schools, 33*, 163-175. https://doi.org/10.1002 /(SICI)1520-6807(199604)33:2<163: AID-PITS10> 3.0.CO; 2-C

Pajares, F., & Valiante, G. (1996). *Predictive utility and causal influence of the writing self-efficacy beliefs of elementary students* [Paper presentation]. The Annual Meeting of the American Educational Association, New York, NY. https://doi.org/10.1080/00220671.1997.10544593

Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology, 24*, 390-405. https://doi.org/10.1006/ceps.1998.0995

Prickel, D. (1994). *The development and validation of a writing self-efficacy scale for adult basic writers and its use in correlational analysis* [Unpublished doctoral dissertation]. Oregon State University. https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/8w32r8160

Qader, R. O., & Yalcin Arslan, F. (2019). The effect of flipped classroom instruction in writing: A case study with Iraqi EFL learners. *Teaching English with Technology*, *19*(1), 36-55.

Raimes, A. (1983). *Techniques in teaching writing*. Oxford University Press.

Rattanadilok Na Phuket, P., & Othman, N. B. (2015). Understanding EFL students' errors in writing. *Journal of Education and Practice, 6*(32), 99-106. http://iiste.org/Journals/index.php/JEP

Ruffini, M. F. (2012). Screencasting to engage learning. *Educause Review Online*.

http://www.educause.edu/ero/article/screencasting-engage-learning

Samiee Zafarghandi, M. (2018). The effect of flip learning on students' self-Efficacy and academic achievement. *Anthropology of Education eJournal*. http://dx.doi.org/10.2139/ssrn.3154001

Schunk, D. H. (1985). Self-efficacy and academic motivation. *Educational Psychologist, 26*(34), 207-231. http://www3.interscience.wiley.com

Sesrica, M., & Jismulatif, A. (2017). The effect of using pictures in teaching writing descriptive text for the second year students of SMPN 1 Siak Hulu. https://media.neliti.com/media/publications/200172none.pdf

Shah P. M., Mahmud, W. H., Din R., Yusof, A., & Pardi, K. M. (2011). Self-Efficacy in the writing of Malaysian ESL learners. *World Applied Sciences Journal (Innovation and Pedagogy for Lifelong Learning)*, *15,* 8-11. http://www.idosi.org/wasj/wasj15(IPLL)11/2.pdf

Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attributions, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement level differences. *Journal of Educational Psychology*, *87*, 386-398. https://doi.org/ 10.1037/0022-0663.87.3.386

Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology*, *81*, 91-100. https://doi.org/10.1037/0022-0663.81.1.91

Sirakaya, D., & Ozdemir, S. (2018). The effect of flipped classroom model on academic achievement, self-directed learning readiness, motivation, and retention. *Malaysian Online Journal of Educational Technology - MOJEST*, *6*(1), 76-86. http://www.mojet.net

Soliman, N. A. (2016). Teaching English for academic purposes via the flipped learning approach. *Procedia - Social and Behavioral Sciences 232*, 122-129. https://doi.org/10.1016/j.sbspro.2016.10.036

Song, Y., & Kapur, M. (2017). How to flip the classroom – "productive failure or traditional flipped classroom" pedagogical design? *Journal of Educational Technology & Society*, *20*(1), 292- 305. http://www.ifets.info

Stannard, R. (2008). *Screen capture software for feedback in language education.* [Paper presentation]. The Second International Wireless Ready Symposium, NUCB Graduate School, Nagoya. https://curve.coventry.ac.uk

Styati, E. W. (2016). Effect of YouTube videos and pictures' effect on EFL students' writing performance. *Dinamika Ilmu, 16*(2), 307-317. https://doi.org/10.21093/di.v16i2.534

Sugar, W., Brown, A., & Luterbach, K. (2010). Examining the anatomy of a screencast: Uncovering common elements and instructional strategies. *International Review of Research in Open and Distance Learning, 11*(3), 1-20. https://doi.org/10.19173/irrodl.v11i3.851

Su Ping, R. L., Verezub, E., Adi Badiozaman, I. F. B., & Chen, W. S. (2020). Tracing EFL students' flipped classroom journey in a writing class: Lessons from Malaysia. *Innovations in Education and Teaching International*, *57*(3), 305-316. https://doi.org/10.1080/14703297.2019.1574597

Szparagowski, R. (2014). The effectiveness of the flipped classroom. *Honors Projects,127,* 1-31. https://scholarworks.bgsu.edu/honorsprojects/127

Tajallizadeh Khob, M., & Rabi, A. (2014). Meaning-focused audiovisual feedback and EFL writing motivation. *Journal of Language and Translation*, *4*(2), 61-70. https://www.sid.ir/en/journal/ViewPaper.aspx?id=513515

Thompson, R., & Lee, M. J. (2012). Talking with students through screencasting:

Experimentations with video feedback to improve student learning. *The Journal of Interactive Technology and Pedagogy, 1*.
https://jitp.commons.gc.cuny.edu/category/issues/issue-one/

Tuna, G. (2017). *An action study on college students' EFL writing skills development through flipped classroom environments* [Unpublished master's thesis]. Middle East Technical University.
http://etd.lib.metu.edu.tr/upload/12621730/index.pdf
https://hdl.handle.net/11511/27053

Unakorn, P., & Klongkratoke, U. (2015). *Effectiveness of flipped classroom to mathematics learning of grade 11 students* [Paper presentation]. The 21st & 22nd International Conference on Language, Education, and Humanities & Innovation.
https://icsai.org/procarch/1iclehi/1iclehi-44.pdf

Vaezi, R., Afghari, A., & Lotfi, A. (2018). Flipped teaching: Iranian students' and teachers' perceptions. *Applied Research on Language Teaching, 8*(1), 139-154.
https://doi.org/10.22108/are.2019.114131.1382

Van Heerden, M., Clarence, S., & Bharuthram, S. (2017). What lies beneath: Exploring the deeper purposes of feedback on student writing through considering disciplinary knowledge and knowers. *Assessment & Evaluation in Higher Education*, *42*(6), 967-977.
https://doi.org/10.1080/02602938.2016.1212985

Villanueva, J. (2016). *Flipped classroom: An action research* [Paper presentation]. The 21st Annual Technology, Colleges and Community Worldwide Online Conference.
http://hdl.handle.net/10125/40822

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Watcharapunyawong, S., & Usaha, S. (2013). Thai EFL students' writing errors in different text types: The interference of the first language. *English Language Teaching, 6*(1),67-78.
http://dx.doi.org/10.5539/elt.v6n1p67

Wening, R. H. (2016, June). The role of picture series in improving students' writing ability [Paper presentation]. *International Conference on Education*, 739-746.

Wilson, S. G. (2013). The flipped class: A method to address the challenges of an undergraduate statistics course. *Teaching of Psychology*, *40*(3), 193-199.
https://doi.org/10.1177/0098628313487461

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, *39*(4), 510-522.
http://www.elsevier.com

Zheng, M., Chu, C., Wu, Y., & Gou, W. (2018). The mapping of on-line learning to flipped classroom: Small private online course. *Sustainability*, *10,* 1-14.
https://doi.org/10.3390/su10030748

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1), 51-59.
https://doi.org/10.1037/00220663.82.1.51

Appendix A
The Analytic Rubric (Jacobs et al., 1981)

Category	Score	Criteria
CONTENT	25-21	**EXCELLENT TO VERY GOOD:** knowledgeable • substantive •thorough development of thesis • relevant to assigned topic
	20-16	**GOOD TO AVERAGE:** some knowledge of subject • adequate range • limited development of thesis • mostly relevant to topic, but lacks detail
	15-11	**FAIR TO POOR:** limited knowledge of subject • little substance •inadequate development of topic
	10-0	**VERY POOR:** does not show knowledge of subject • non-substantive • non pertinent • OR not enough to evaluate
ORGANIZATION	25-21	**EXCELLENT TO VERY GOOD:** fluent expression • ideas clearly stated/ supported • succinct • well-organized • logical sequencing • cohesive
	20-16	**GOOD TO AVERAGE:** somewhat choppy • loosely organized but main ideas stand out • limited support • logical but incomplete sequencing
	15-11	**FAIR TO POOR:** non-fluent • ideas confused or disconnected • lacks logical sequencing and development
	10-0	**VERY POOR:** does not communicate • no organization • OR not enough to evaluate
LANGUAGE USE	25-21	**EXCELLENT TO VERY GOOD:** effective complex constructions • few errors of agreement, tense, number, word order/function, article, pronouns, prepositions
	20-16	**GOOD TO AVERAGE:** effective but simple constructions • minor problems in complex constructions • several errors of agreement, tense, number, word order/function, article, pronouns, prepositions but meaning seldom obscured
	15-11	**FAIR TO POOR:** major problems in simple/ complex constructions • frequent errors of negation, tense, number, word order/function, article, pronouns, prepositions and/ or fragments, run-ons, deletions • meaning confused or obscured
	10-0	**VERY POOR:** virtually no mastery of sentence construction rules • dominated by errors • does not communicate • OR not enough to evaluate
VOCABULARY	15-13	**EXCELLENT TO VERY GOOD:** sophisticated range •effective word/idiom choice and usage • word for mastery • appropriate register
	12-10	**GOOD TO AVERAGE:** adequate range • occasional errors of effective word/idiom form, choice, usage but meaning not obscured
	9-7	**FAIR TO POOR:** limited range • frequent errors of effective word/idiom form, choice, usage • meaning confused or obscured
	6-0	**VERY POOR:** essentially translation • little knowledge of English vocabulary, idioms, word form • OR not enough to evaluate
MECHANICS	10	**EXCELLENT TO VERY GOOD:** demonstrates mastery of conventions • few errors of spelling, punctuation, capitalization, paragraphing
	9-8	**GOOD TO AVERAGE:** occasional errors of spelling, punctuation, capitalization, paragraphing but meaning not obscured
	7-6	**FAIR TO POOR:** frequent errors of spelling, punctuation, capitalization, paragraphing • poor handwriting • meaning confused or obscured
	5-0	**VERY POOR:** no mastery of conventions • dominated by errors of spelling, punctuation, capitalization, paragraphing • handwriting illegible • OR not enough to evaluate

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

The Effect of Implementing Flipped Classroom Model on Critical Thinking and Performance of Iranian EFL Learners in Learning Grammar

Research Article
pp. 183-204

Zahra Afzali[1]
Siros Izadpanah*[2]

Received: 2020/09/06 Accepted: 2021/01/02

Abstract
The flipped classroom is a teaching approach that can afford additional face-to-face communication with students in the classroom. Although some studies have already investigated the flipped classroom in general, few studies have investigated critical thinking and performance in learning grammar, in particular. The design of the study was quasi-experimental. This research was performed within an English class for six weeks in 2020. Three hundred and sixty English learners were chosen through multistage cluster sampling from two different institutions within the age range of 13-19 who were studying at intermediate and upper-intermediate levels. Eight grammar contents were selected among fifty grammar videos based on the content validity index (CVI) and the content validity ratio (CVR). EPT (English performance tests) and CT (critical thinking) questionnaires were used. The experimental group was educated with the flipped classroom model, whereas the courses were performed based on the present syllabus in the control group. It was realized that there was no substantial discrepancy between the pre-test and post-test scores of the control group, while there was a significant difference between the pre-test and post-test scores of the experimental group. Teachers can be proposed to manipulate the flipped classroom model to boost learners' critical thinking and performance.

keywords: critical thinking, flipped classroom, flipped learning, flipped teaching, performance

*Corresponding Author
[1] MA Student, Department of English Language Teaching, Zanjan Branch, Islamic Azad University, Zanjan, Iran. afzalizahra90@gmail.com
[2] Assistant Professor, Department of English Language Teaching, Zanjan Branch, Islamic Azad University, Zanjan, Iran. sirosizadpanah@iauz.ac.ir

DOI: 10.22051/LGHOR.2021.32997.1362

Introduction

The flipped classroom was first mooted as an instructional strategy by Baker in the late 1980s (Segolsson et al., 2017). The educational world has changed its perspective from teaching-centered into learning-centered (Jungst et al., 2003). This disposition shows that learning assumes a key role in every single academic choice (Lombardini et al., 2018). One of the most critical difficulties in the present instructive world is how we can prepare learners to learn well (Wang, 2020). Flipped learning denotes an instructional technique that draws in learners in their learning procedure through cooperative and issue-based learning exercises to build up their basic reasoning and critical thinking aptitudes (Zou, 2020). Despite the fact that teacher-student contact hours do not change, learners can see the instructional materials at home and have more teacher-student commitment in the class. The main goal is to move the instructor-focused on guidance to student-focused on figuring out how to change the educator's work. There is a lack of investigation concentrating on this issue (Hamdani, 2019). To the best of the researchers' knowledge, few studies have examined issue. Traditional instructional techniques have turned out to be insufficient in gathering learning applications. The reason is that students don't feel good and the learning condition isn't proper for them in the conventional guidance framework where the instructor is the focal point of the learning procedure (Cooper, 2014). Students are inactive learners in customary guidance frameworks, making learners careless during the learning process. In such a manner, imaginative learning methodologies can be utilized. Without a doubt, paying more heed to innovation in the learning process will give learners modern skills and rectify the instruction framework by giving genuine training modification (Andujar et al., 2020).

The study on the flipped classroom approach in English institutions is still incomplete, especially research conducted in an Iranian context. Nevertheless, the flipped classroom relates well to the Iranian curriculum and how teachers are supposed to work within the communicative classroom (Shahamat et al., 2019). This main characteristic makes learners study and practice hard, in which the flipped classroom is the basic savior to help them.

This study is to research the effects of the flipped classroom on learners' critical thinking and performance in learning grammar for Iranian intermediate and upper-intermediate EFL learners.

The study investigated the following research questions:

Q1. Does the experience of using a flipped-classroom approach significantly improve upper-intermediate learners' grammatical performance?

Q2. Does the experience of using a flipped-classroom approach significantly improve upper-intermediate learners' critical thinking?

Q3. Does the experience of using a flipped-classroom approach significantly improve intermediate learners' grammatical performance?

Q4. Does the experience of using a flipped-classroom approach significantly improve intermediate learners' critical thinking?

Literature Review

Flipped Classroom

It was first presented as an instructional strategy by Baker in the late 1980s. On account of the advances of personal computer(PC) development, the Internet,

and YouTube in 2006, Baker's considerations have been recognized and framed into what today is known as the flipped classroom. The flipped classroom is an instructional approach that gives the students different opportunities to process instructional materials (Han et al., 2020; Judy & Huang, 2020). In a flipped classroom, learners utilize some sources provided by their teacher prior to classroom presentation, while classroom time is mainly dedicated to group cooperation and problem-solving activities.

The expectation behind the flipped classroom is to keep the close and personal time between the teacher and students so the activity can be used to process the material rather than as in typical talks where the students are latent beneficiaries of data (Awidi, & Paynter, 2019; Bond, 2020; Hamdani, 2019; Hung, 2015; Nicolosi, 2014). The flipped classroom has the same formation as an Instructive classroom. Consequently, the four needs of the communicative classroom are closely related to how one is supposed to work with the flipped classroom method. For instance, as indicated by Chan et al. (2020) and Van Alten et al. (2019), the students and teacher have greater opportunities to use and contextualize the target language in different collaborative learning exercises to highlight the different social uses of language as more time is freed in the classroom.

Performance

Performance is the result and outcome of an individual's effort in an official learning environment that is the main goal of each pedagogical system. It is one of the most important criteria. Performance means the sequence of responses that aim to alternate the environment with definite solutions. The sequence of responses means the respondent's behavior which leads to production (Henderson et al., 2020; Özkurkudis & Bümen, 2019). While the theoretical description of the performance has been provided, the practical definition of this attribution in this study is still vague. Performance, refers to the obtained scores in the final exam in grammar sections, which were obtained after one term of experiencing the flipped classroom environment.

Critical Thinking

Critical thinking and analyzing ideas have always been discussed since ancient times and back to the Greek philosopher era. As it has been mentioned, the origin of critical thinking goes back to Aristotle, Socrates, and Plato's time. While Socrates focuses on subjectivism and the willingness to criticize, which is required for human achievement, Aristotle and Plato worked on the agenda of the significance of logic and critical thinking in achieving freedom (Chance,1986).

Science and technology have progressed in the 21st century (Andujar et al., 2020). The skills of the 21stcentury involve countless skills like critical thinking, association, and innovation (Andujar et al., 2020). However, the main thinking skill that is important in the 21st century is critical thinking (Ahadiat, 2019). It is a skill that is required to enhance the thinking ability of the learners (Dooly & Sadler, 2020). Also, it is essential for problem-solving. Taylor et al. (2020) believe that, to solve problems, we need to think deeply to evaluate the problems through the use of existing knowledge and based on critical thinking.

One characteristic of critical thinking is that it starts with comprehending our thinking (Boss & Larmer, 2013; Chen et al., 2014; Elmaadaway, 2018). According to these researchers, we cannot transfer non-propositional knowledge just

by the language, and via speaking to learners; they should learn it practically. It means that how to think is different from what to think about. In utilizing and producing a new language, using its structures and grammar are crucial factors. Comprehending and knowing these structures requires deep and critical thinking to evaluate the problem through the use of grammatical knowledge.

Method

Participants

The participants of this research were 360 English learners aged 13-19 who were studying at intermediate and upper-intermediate levels from two institutions in Zanjan, Iran. They were selected through multiple-stage cluster sampling. Out of 110 institutions in Zanjan province, Iran, researchers selected two institutions. Nearly 30% of students were at intermediate and upper-intermediate levels (N = 5830). Out of this number, about 3485 learners were females and about 2345 were males. Based on Morgan's table, 360 learners were required for the current report (210 females, 150 males). The institutions included the Science and Technology of Kish and the Iran Language Institute (ILI). The exact number of students in the institutions mentioned was 370 in the summer of 2020.

Two classes at the intermediate level and two upper-intermediate classes (based on a placement test) were selected from each institution. As a whole, there were 8 classes, 4 male classes (n = 146) and 4 female classes (214). The Oxford placement test was conducted for the 8 classes as a pretest to have a homogeneous group. Group A was identified as upper-intermediate level students (66 males and 94 females) and Group B as intermediate level students (80 males and 120 females). Students who were able to answer 28-36 questions out of 60 test items were regarded as an intermediate level and those correctly responding 48-55 questions were assigned to the upper-intermediate proficiency level. The learners in the experimental group were taught grammatical points and strategies through flipped classroom by their teacher. Thus, by collaborating with the researcher, the teacher played the role of a facilitator. The participants in the experimental group got familiar with such skills and employed them to enhance their grammatical ability. In contrast, the control group learners were taught the course using traditional approaches for all lessons.

Table 1

Estimation of Sample Size

Z	Confidence Level	Relative Error D	Sample Size N	Volume of Society N
1.96	0.95	0.05	360	5830

Instruments

Oxford Placement Test (OPT). The language proficiency test version 2 including 60 items (such as matching, comprehension text, and multiple-choice questions) was conducted to reach the previously mentioned objectives to confirm the students' similarities. Thirty minutes were offered to the students to reply to the survey questions that were concentrated mostly on reading abilities, grammar, and vocabulary.

Ricketts' (2003) Critical Thinking Dispositions Questionnaire. It was used to measure the intermediate and upper-intermediate EFL learners' critical

thinking disposition. To achieve the goals mentioned before, a test was considered to get feedback from learners. In addition, that feedback was about skills like reading and grammar, etc. The time of the test was about 30 minutes. In the test, there were 33 statements; the lowest and the highest scores were 33 and 145 and the average score was 99. There were 3 subscales of 1. Innovation —> 11 statements 2. Maturity —> 9 statements 3. Engagement —> 13 statements. Each subscale has a coefficient, respectively: Innovation = 0.64, maturity = 0.53, engagement = 0.82, and the coefficient of the instrument was 0.76. The reliability coefficient of the instrument had been reported to be about 0.76 by Pakmehr et al. (2013).

EPT (English Performance Test). In this study for evaluating students' performance, EPT was used. It is validated by Dortaj (2004) for Iranian society. It had 47 questions on a 5 Likert scale. Those with lower than 120 were considered low educational performance, those with more than 175 high performances, and those with a score between 121-174 were considered medium performance. It assesses performance five domains, self-efficacy, emotional impacts, lack of outcome control, time schedule, and motivation. After extracting the data from the questionnaires, the Cronbach's Alpha value of the variables was calculated as follows:

Table 2
The Result of the Cronbach's Alpha Test

Questionnaire	Question number	Cronbach's Alpha
Performance	48	0.981
Critical thinking	33	0.975

Grammar Videos. There were fifty different types of grammar subjects. The researchers chose 8 important topics. Those topics were asked by experts and experienced teachers based on CVR and CVI formulas. Lawsh formula is as follows:

$$CVR = \frac{(Ne - N \div 2)}{N \div 2}$$

In this study, students accessed video grammars at home and did exercises in the flipped classroom with teachers' guidance and supervision. Learners watched the grammatical movies at home and the next session teacher taught those contents in the class and they did the exercises and appropriate pair works. It was done for six weeks. Control groups were taught through traditional methods.

Procedures

To administer this research following steps were taken:

Administering the Placement Test. To begin the study, grammar was chosen for eight classes at the Iran Language Institute (ILI) and the Kish Institute of Science and Technology. To have a homogeneous group, the Oxford placement test was conducted for the 8 classes as a pretest. Group A was identified as upper-intermediate level students (66 males and 94 females) and Group B as intermediate level students (80 males and 120 females). Students who were able to answer 28-36 questions out of 60 test items were regarded as an intermediate level and those correctly responding 48-55 questions were assigned to the upper-intermediate proficiency level.

Administering the Pre-test for Group A. In the next stage, some films that have been gathered by researchers based on the upper intermediate course book's content were used. Those films were about the grammatical points including causative structure, infinitives, gerunds, relative pronouns, future perfect, progressive, and phrasal verbs that the researchers gave to the participants. They were asked to watch the film before the teaching session to be argumentative, have more interaction among students and enhance their performance during grammar exercises.

Administering the Pre-test for Group B. In this group, like the previous one, they were given another group of grammatical movies to be watched at home before teaching sessions. Those grammars were chosen based upon the coursebook content for intermediate learners. They were about tenses, conditional sentences, passive voices, and part of speech. Instructors asked students to watch the movies in advance. Afterward, some questions and ambiguities have been provoked in their mind that needs to be solved. Before teaching in the classroom, this background information leads to more critical thinkers and can ask more controversial questions and improve their performance during the session.

Design of the Study

The design of the study was quasi-experimental research. Participants were randomly allocated to two groups; however, this research had an experimental evaluation group design since there was a control group. The key independent variable was the role of a flipped classroom with two levels (intermediate vs. upper-intermediate), and the dependent variables included two dimensions of critical thinking and grammatical performance.

Data Analysis

To examine the data gathered from these applicants, SPSS 22 software was manipulated. A quasi-experimental study was assigned in order to study the effect of the flipped classrooms on learners' critical thinking and performance in learning grammar for Iranian intermediate and upper-intermediate level learners.

Results

Descriptive statistics for the upper-intermediate experimental group was calculated and the outcomes were given in the following:

Table 3
Central Tendencies and Dispersion for Variables in Experimental and Control Groups for Upper-Intermediate Learner (Descriptive Statistics Related to The Variables of Hypotheses 1 and 2 of the Research)

		Mean	Std. Deviation	Minimum	Maximum
Experimental Group	Performance pre	104.16	11.07	92.00	125.00
	Performance post	255.11	28.60	217.00	301.00
	Critical thinking pre	84.28	15.82	65.00	126.00
	Critical thinking post	121.54	10.10	102.00	135.00
	Valid N (listwise)				
Control Group	Performance pre	108.38	7.68	96.00	119.00
	Performance post	157.90	11.98	49.00	257.00
	Critical thinking pre	90.89	9.84	73.00	106.00
	Critical thinking post	88.70	7.90	76.00	105.00
	Valid N (list wise)				

Table 3 shows the central indicators and the dispersion of research in upper-intermediate learners. Mean performance and critical thinking in the experimental group's post-test significantly increased.

Table 4
Central Tendencies and Dispersion for Variables in Experimental & Control Groups for Intermediate Learners (Descriptive Statistics Related to the Variables of Hypotheses 3 and 4 of the Research)

Group		Mean	Std. Deviation	Minimum	Maximum
Experimental Group	Performance pre	114.63	22.05	101.00	198.00
	Performance post	188.92	13.44	170.00	214.00
	Critical thinking pre	71.90	8.86	60.00	89.00
	Critical thinking post	136.14	10.67	119.00	151.00
	Valid N (list wise)				
Control Group	Performance pre	104.04	10.31	77.00	117.00
	Performance post	112.92	12.72	83.00	137.00
	Critical thinking pre	71.13	8.70	60.00	92.00
	Critical thinking post	79.02	7.99	65.00	94.00
	Valid N (list wise)				

Table 4 shows the central indicators and the dispersion of research variables in intermediate learners. Mean performance and critical thinking in the experimental group's post-test significantly increased.

Inferential Analysis

In this study, the Kolmogorov - Smirnov test was utilized to test the hypothesis of normality of the scores.

Table 5

Test the Hypothesis of Normalization of Research Variables at a Higher Average Level in Research Hypotheses 1 and 2)

Group	Variable	Kolmogorov-Smirnov Z	Asymp. Sig. (2-tailed)	Test result
Experimental Group	Performance pre	1.11	0.12	Acceptance of the H_0
	Performance post	1.17	0.11	Acceptance of the H_0
	Critical thinking pre	1.15	0.94	Acceptance of the H_0
	Critical thinking post	1.15	0.12	Acceptance of the H_0
Control Group	Performance pre	1.28	0.08	Acceptance of the H_0
	Performance post	1.10	0.21	Acceptance of the H_0
	Critical thinking pre	1.23	0.10	Acceptance of the H_0
	Critical thinking post	1.26	0.09	Acceptance of the H_0

The results of Table 5 show that the assumption of normality of research variables for performing the ANCOVA test in hypotheses 1 and 2 is established ($P < 0.05$).

The results of Table 5 show that the assumption of normality of research variables for performing the ANCOVA test in hypotheses 1 and 2 is established ($P < 0.05$).

Table 6
One-Sample Kolmogorov-Smirnov Test to Test the Hypothesis of Normalization of Research Variables at the Intermediate Level in Research Hypotheses 3 and 4

	Variable	Kolmogorov-Smirnov Z	Asymp. Sig. (2-tailed)	Test result
Experimental Group	Performance pre	1.05	0.20	Acceptance of the H_0
	Performance post	1.23	0.10	Acceptance of the H_0
	Critical thinking pre	1.52	0.05	Acceptance of the H_0
	Critical thinking post	1.122	0.12	Acceptance of the H_0
Control Group	Performance pre	1.12	0.18	Acceptance of the H_0
	Performance post	0.75	0.65	Acceptance of the H_0
	Critical thinking pre	0.52	0.73	Acceptance of the H_0
	Critical thinking post	0.83	0.55	Acceptance of the H_0

The results of Table 6 show that the assumption of normality of research variables for performing the ANCOVA test in hypotheses 3 and 4 is established (P>0.05).

Homogeneity of Variances

In this research, the Levene test was applied to investigate the variance of the variance.

Table 7
Homogeneity of the Variances in Hypotheses 1 and 2

	Levene Statistic	df1	df2	Sig.
Performance pre	1.69	1	158	0.08
Performance post	1.42	1	158	0.08
Critical thinking pre	1.07	1	158	0.21
Critical thinking post	1.85	1	158	0.07

The results of Table 7 show that the assumption of homogeneity of variance of research variables to perform the ANCOVA test in hypotheses 1 and 2 is established (P < 0.05).

Table 8
Homogeneity of Variances of Intermediate Learners in the Pretest (Examining the Default in Hypotheses 3 and 4)

	Levene Statistics	df1	df2	Sig.
Performance pre	2.62	1	198	0.11
Performance post	0.17	1	198	0.69
Critical thinking pre	1.03	1	198	0.31
Critical thinking post	1.93	1	198	0.23

The results of Table 8 show that the assumption of homogeneity of variance of research variables for the ANCOVA test is set in hypotheses 3 and4 ($P < 0.05$).

Covariate Run Before Starting the Research

This hypothesis was met before the independent variable and the flipped learning method was implemented.

Homogeneity of Regression Slope

To check the homogeneity of the regression slope, researchers needed to compute the F value of the communication between the covariate and the independent variable. The results of the study were as follows:

Table 9

Regression Slope Homogeneity Test for Upper Intermediate Level in Hypotheses 1 and 2

Source	F	Sig.	Conclusion
group * Performance pre	1.04	0.35	Verify regression slope homogeneity
group* Critical thinking pre	1.61	0.28	Verify regression slope homogeneity

The results of Table 9 show that it is accepted in examining hypotheses 1 and 2($P < 0.05$).

Table 10

Regression Slope Homogeneity for Intermediate Level in Hypothesis 3 and 4

Source	F	Sig.	Conclusion
group*Performance pre	2.66	0.09	Verify regression slope homogeneity
group*Critical thinking pre	0.97	0.56	Verify regression slope homogeneity

The results of Table 10 show the homogeneity of the regression slope of the research is established in examining hypotheses 3 and 4 ($P > 0.05$).

The Linearity of the Correlation

To check the linearity of the correlation of the independent variable and the covariate, researchers needed to estimate the F value of the correlation variable. The results of the survey are as follows:

Table 11

The Independent and the Dependent Variable in the Upper-Intermediate Group in Hypothesis 1 & 2

Source	F	Sig.	Conclusion
Performance pre	36.43	0.00	Linear correlation between the covariate and the independent variable
Critical thinking pre	30.15	0.00	Linear correlation between the covariate and the independent variable

The results of Table 11 show that the correlation of the independent variable and the pretest is accepted in hypotheses 1 and 2 ($P < 0.05$).

Table 12
The Linearity of the Correlation of the Independent and the Dependent Variable in the Intermediate Group in Hypotheses 3 and 4

Source	F	Sig.	Conclusion
Performance pre	73.22	0.00	Linear correlation between the covariate and the independent variable
Critical thinking pre	50.93	0.00	Linear correlation between the covariate and the independent variable

The results of Table 12 show that the correlation of the independent variable and the pretest are accepted in hypotheses 3 and 4($P < 0.05$).

Testing Research Hypotheses

Hypothesis 1: Flipped Learning Improves Upper-Intermediate Learners' Grammatical Performance. Covariance analysis was run to investigate hypothesis one. The outcomes of the analysis of covariance were listed in the following:

Table 13
The Result of the Analysis of Covariance for Upper-Intermediate Students' Grammatical Performance, the Result of the Analysis of Covariance for Upper-Intermediate Students' Grammatical Performance

Group			Mean	Std. Deviation	F	Sig.
Upper Learners	Performance pre	experiment	104.16	11.07	197.77	0.00
		control	108.38	7.68		
	Performance post	experiment	255.11	28.60		
		control	157.90	11.98		

The results of covariance analysis 13 show that flipped Learning improves the grammatical performance of upper intermediate language learners ($F = 197.77$, $P < 0.01$) and it could be concluded that the mean of the two groups at post-test after adjusting for pre-test scores were significantly different. As shown in the above table, the mean grammatical performance of upper-intermediate learners' in the control group was 108.38 in the pre-test, and 157.90 in the post-test, while the average grammatical performance of upper-intermediate learners' in the pre-test group was 104.16 and in the post-test was 255.11. Thus, it was determined that by excluding the pre-test effect, flip learning improves the grammatical performance of upper-intermediate learners. Based on the above explanation, hypothesis 1 is accepted.

Hypothesis 2: Flipped Learning Improves Upper-Intermediate Learners' Critical Thinking. Covariance analysis was used to test the above hypothesis. The outcomes of the analysis of covariance were listed in the following:

Table 14
The Result of the Analysis of Covariance for Upper-Intermediate Students' Critical Thinking

Group			Mean	Std. Deviation	F	Sig.
Upper-intermediate Learners	Critical thinking pre	experiment	84.28	15.82	651.23	0.00
		control	90.89	9.84		
	Critical thinking post	experiment	121.54	10.10		
		control	88.70	7.90		

Results of analysis of covariance Table 14 show that flipped Learning improves the critical thinking of upper intermediate level learners (F = 651.23, P < 0.01). As seen in the above table, in terms of critical thinking, in the pre-test, the mean scores of upper-intermediate learners in the experimental and control groups were 90.88 and 88.70, respectively, while in the post-test, the mean score of upper-intermediate learners in the control group was 84.28 and in the experimental group was 121.54. Taking into account, it was determined that by deleting the pre-test effect, the flipped learning method could improve upper-intermediate learners' critical thinking. Based on the above explanation, hypothesis 1 is accepted.

Hypothesis 3: Flipped Learning Improves Intermediate Learners' Grammatical Performance. Covariance analysis was used to test the above hypothesis. The outcomes of the analysis of covariance are listed in the following tables:

Table 15
The Result of the Analysis of Covariance for Intermediate Learners' Performance

Group			Mean	Std. Deviation	F	Sig
Intermediate learners	Performance Pre	experiment	114.63	22.05	1187.27	0.00
		control	104.04	10.31		
	Performance post	experiment	188.92	13.44		
		control	112.92	12.72		

The results of the analysis of covariance in Table 15 show that flipped Learning improves the performance of intermediate language learners (F = 1187.27, P < 0.01). Thus, it could be concluded that the mean of the two groups at post-test after adjusting for pre-test scores were significantly different. The mean grammatical performance scores of intermediate level learners in the pre-test and control groups were 104.04 and 112.92, respectively, while the mean grammatical performance scores of intermediate learners in the pre-test and control groups were 114.63 and 188.92, respectively. It was deduced that flip learning improves intermediate-level learners' grammatical performance. Based on the above explanation, hypothesis 1 is accepted.

Table 16
The Outcome of the Analysis of Covariance for the Intermediate Learners' Critical Thinking

Group			Mean	Std. Deviation	F	Sig
Intermediate learners	Critical thinking pre	experiment	71.90	8.86	2263.90	0.00
		control	71.13	8.70		
	Critical thinking post	experiment	136.14	10.67		
		control	79.02	7.99		

Hypothesis 4: Flipped Learning Improves Intermediate Learners' Critical Thinking. The results of analysis of covariance in Table 16 show that flipped learning improves the critical thinking of intermediate-level learners (F = 2263.90, P < 0.01). As displayed in the table, the value of F in the covariance analysis for the critical thinking ability of intermediate level students were found to be 2263.90 and since Sig = 0.000 was less than 0.05, it was significant at the 0.05 level. Therefore, it could be concluded that the mean scores of the two groups in the post-test, after adjusting the pre-test scores, was significantly different. As displayed

in the above table, regarding critical thinking, the mean score of intermediate level learners in the control group was 71.13 in the pretest and 79.02 in the posttest while the mean score of the intermediate level learners in the experimental group was 71.90 in the pretest and 136.14 in the posttest, it was accepted that flipped learning method significantly improves critical thinking for the intermediate level students. Based on the above explanation, hypothesis 1 is accepted.

Discussion

The first research question was, "Does the experience of using a flipped classroom approach significantly improve upper – intermediate learners' grammatical performance in learning grammar?" According to the results of the covariance analysis, it was found that the flipped classroom improves upper-intermediate learners' grammatical performance. Taking into consideration, it was determined that by excluding the pre-test effect, flip learning improves the grammatical performance of upper-intermediate learners. The findings of this study are similar to the study by Luo et al. (2017). Their empirical outcomes showed that learner agency had an important influence on learners' academic performance. Outcomes in this study show that performance will increase in the flipped classroom. Another investigation was carried out by Wang (2020), matching the impacts of individual vs. group face-to-face class activities in the flipped classroom on learner's test performances. This study discussed the performance through individual and group face-to-face performances in the flipped classroom. The research showed that the mean test score of learners in face-to-face class performances was more effective than individual face-to-face class performances and the difference across Group 1 and Group 2 was considerable The flipped classroom model increases the area knowledge of the learners (Hung, 2015) and learners' success (Baepler et al., 2014; Chen et al., 2014; Liebert et al., 2015; Mclaughlin & Rhoney, 2015).

Learners actively take part in the process and their participation in the lessons increases (Nicolosi, 2014; Özkurkudis, & Bümen, 2019; Roach, 2014). Therefore, it can be considered that the achievement obtained in the group where the flipped classroom model was implemented, is due to the aforementioned reasons. The increase in the active participation of students in the lesson is reflected favorably in their performance. This study will be useful for teachers, curriculum designers, and academic departments.

The second research question was, "Does the experience of using a flipped-classroom approach significantly improve upper – intermediate learners' critical thinking in learning grammar?" The methods used to improve student academic achievement and critical thinking skills in learning have been discussed seriously in the education environment for the past two decades (Washaw & Openshaw, 2011).

The flipped classroom is an important part of general education (Bergmann & Sams, 2012, Sophia Learning, 2013). Flipped classroom effect on student academic achievement and student critical thinking skills in the high school learning classroom has not been suggested in articles, journals, or research studies on this new approach.

As discussed before, Ricket's (CDTI) questionnaire was used for measuring critical thinking, which includes 33 Likert scale alternatives. Based upon the results of covariance analysis, the flipped classroom was found to improve

upper- intermediate learners' critical thinking. As was revealed in the results, the value of F in the covariance analysis for the critical thinking ability of upper-level students was 651.230. Since sig = 0.000 was less than 0.05, it was significant at the 0.05 level as relates to critical thinking, in the pretest, the mean scores of upper-intermediate learners in the experimental and control groups were 90.88 and 88.70, respectively, while in the post-test, their mean scores were 84.28 and 121.54, respectively implying that critical thinking ability is significantly affected by the flipped classroom environment.

The findings of this research questionnaire is in line with that of Yazici et al. (2020) examined the correlation between critical thinking skills and educational performance and probed the degree to which demographic characteristics could moderate the relationship between the variables of the study. They found that there is an important correlation between them and explored that there are differences across cultures. Also, the results are in acceptance with that of Evseeva, and Solozhenko, (2015) who sought to find out if there was any relationship among self-efficacy, thinking styles, critical thinking, emotional intelligence, and academic achievement in nursing students. They found that the variables of the study are related to academic success. Winch (2020) also investigated the impact of flipped learning on nursery students' critical thinking. In sync with our findings, the results showed some satisfaction in the learning and crucial improvement in the critical thinking aptitude of the learners.

Although the impacts and whether they directly affect students' grammatical progress cannot be accurately described, this current study shows that increased success seems to be determined by three aspects: pre-class activities, in-class activities, and process teaching. The above results clearly show the positive effects on students' grammatical progress in ICT (information and communications technology).

In the first place, the homework such as do a search on media led to many different ways of learning and drew students' attention and change them to be more accountable and self-supporting learners whose activities are dramatically boosting also, stimulate to be taught. The research obtained aligns with Kang's outcome (2015), which showed that homework included video content and diverse teaching programs; considering that students can easily show their interest. Kang's educational tool, a mobile class board, also turns the class atmosphere into a better place.

With more practice and effort in the ICT-aided flipped-classroom, students are more independent and are willing to self-study prior to their classes. Before starting each class training session, learners can answer most of the teacher's questions and get a good score with enough practice. Timely attendance in the classroom ensures regularity and also all students can use the class effectively. Also, the teacher's teaching greatly impacts the student's learning (Evseeva & Solozhenko, 2015; Hung, 2015; Zhonggen & Guifang, 2016).

Secondly, participatory activities in the classroom can satisfy students and increase their motivation to learn. It can also increase their knowledge through tests, educational games, discussion, and writing. This supports Zhonggen and Guifang's (2016) findings of students' satisfaction in the clicker-aided flipped classroom. Denprapat and Chuaychoowong (2016) stated that this increases motivation and it is an opportunity for students to express different abilities and learn different skills.

Eventually, in a not-completely new learning environment, the teaching method was simple, structured, and sufficient for the students to adapt themselves. This current study also verified the results of Denprapat and Chuaychoowong (2016), which suggested that one element of English language enhancement was the teaching process. The teaching procedure was not only simple but also systematic in this current study. As a result, each step the teacher asked them to complete could be understood by the students. Each step was systematically connected from one to another, i.e. after the presentation to practice and later develop. The teaching proceeding by incorporating the flipped model, communicative approach, and technology was neither too conventional nor too new.

English language teachers would benefit from the results of the study as well. They should consider their teaching strategies and methods and find creative ways to enhance students' critical thinking in the flipped classroom to get better results.

The third research question was, "Does the experience of using a flipped classroom approach significantly improve intermediate learners' grammatical performance in learning grammar?"

According to the results from the covariance analysis, a flipped classroom improves intermediate learners' grammatical performance. The value of F for learners' grammatical performance at the intermediate level was 1187.27. Since Sig = 0.000 was less than 0.05, it was significant at $P < 0.05$., the mean scores for the grammatical performance of intermediate level learners were 104.04 and 114.63 for the control and experimental group, respectively, while in the posttest, their mean scores for the grammatical performance were 112.92 and 188.92 for the control and experimental group, respectively. It was assumed that flipped learning improved the grammatical performance of intermediate learners.

All of the students in control and experimental classes scored low on the grammar pretest (62.68 and 65.85 and got an equal score in the pretest. After a while when they join the flipped classes, they could be able to improve their English workmanship. As Farah (2014), many scientists proved that this kind of classes could affect students' scores by providing enough time to learn and repeat English grammar. Students also in such classes have more time to focus on what grammar is about and what concept and application it wants to convey? Other specialists such as Mason et al. (2013); Murphree (2014); Strayer (2012); and Wilson (2013), found that learners will get better by flipped learning.

Surveys about post-tests indicate that adopting the flipped classroom strategy appears beneficial in developing students' grammar knowledge, thus the score of the experimental group was more than that of the control group. This survey was compatible with other cultural and educational contexts such as Kang (2015) and Han (2015), whose studies perceived that the flipped classroom developed students' grammar knowledge compared to other instructional methods. Particularly encountering the unusual challenges of the Saudi context, analyzing students' reactions in this study revealed that flipping English grammar classes could assist in solving some of the non-problematic issues of Saudi secondary schools found by Al-Seghayer (2015) and Rahman and Alhaisoni (2013). The results of the study are consistent with those of Dehghanzadeh and Jafaraghaee (2018), and Uzunboylu and Karagozlu (2015).

The results of Hung's (2015) and Kang's ((2015) studies significantly confirmed and showed the effect of the flipped classroom on the English progress of

first-year students. Similarly, Zhonggen and Guifang (2016), and Denprapat and Chuaychoowong (2016) confirmed the progress of both groups in the post-test and all of these significant improvements have been reflected in the findings of Liu (2016), Denprapat, and Chuaychoowong (2016). This study has demonstrated the possibility of completing the class model to facilitate the English language learners and has confirmed that this sheet model can be used by combining technology at different levels of English language courses.

Flipped learning created an environment for cooperative learning, which allowed what Crouch and Mazur (2001) describe as an opportunity for participants to engage with one another under the instructor's supervision effectively. The pre-class preparation resulted in students gaining the skills to deal with complex challenges and it provided them with some knowledge to share and questions to ask in class.

The participants experienced other benefits of learning grammar through flipped learning. Most students declared that they could retain the knowledge they gained from their grammar course and achieve better course grades. Another effect of flipped learning was related to learning beliefs. The majority of the students believed their role as students changed. They felt they could share knowledge with their teacher and their peers. Moreover, when asked about the possibility of using flipped learning in other courses, more than half of the students confirmed they were able and willing to do so. Finally, the participants were also asked about any signs of autonomous learning. Around half of the students claimed they did not completely rely on their teacher to find additional and related websites. They independently explored the Internet to search for more information related to their grammar course. This is consistent with Amiryousefi's (2019, p. 10) claim that flipped learning can "enhance students' self-study skills and autonomy."

The study shows that students learning in the flipped classroom make themselves more cooperative and raise their motivation and joy in grammar courses. Thus, they can learn faster and it is vivid that they would like to learn more. Also, such findings can be ascribed to the fact that utilizing Flipped classroom with the experimental group raised their interactivity and cooperation and their inspiration to memorize, which increased their satisfaction and adore using the learning language structure.

In addition to increasing the motivation of English learners, the researchers realized both the control group and the experimental group had completely different atmospheres. On the one hand, by accepting that grammar is a difficult subject to learn that requires deep concentration and understanding, the students in the control group showed a lack of interest and boredom. On the other hands, the students in the class with the experimental group experienced learning grammar through the flipped classroom, and the active and positive atmosphere of their classes led to their greater interest and participation. As an inference from the first hypothesis, we conclude that the flipped classroom has a direct and positive effect on students' progress in grammar due to its relaxed atmosphere. These findings from the present study will be useful for teachers, curriculum designers, and academic departments.

The fourth research question was "Does the experience of using a flipped classroom approach significantly improve intermediate learners' critical thinking in learning grammar?"

According to the covariance analysis results, flipped classrooms improve intermediate learners' critical thinking. These findings imply the effectiveness

of flipped classrooms in promoting students' critical thinking and thereby, critical thinking skills. Previous studies also reported that the flipped classroom was associated with positive outcomes. The findings of this research are in line with Coates (2007), Jungić et al. (2015), Kong (2015), and Lai and Hwang (2016), who came up with similar results.

According to a study by Johnson and Renner (2012), different results were obtained in teaching computer science in the flipped classroom. Therefore, according to the results, due to the lack of students' acceptance of the new method and their insistence on traditional methods, it is inferred that this is just an unsuccessful attempt at the flipped classroom teaching method in their research and cannot be applied. As reported by our participants, one of the disadvantages of flipped classrooms was the limited amount of time for teacher's lecturing in the classroom.

Conclusion

Based on the current study findings, the following conclusions were derived from the effect of the flipped classroom on the performance and critical thinking of Iranian EFL learners in learning grammar:

1. The flipped classroom proved more useful and more advantageous than the traditional method for learning English, mainly grammar.
2. Flipped classrooms give students a better opportunity for success and have better efficiency in learning English.
3. A flipped classroom increases the amount of students' learning activities and collaboration.
4. In the flipped classroom, students are more relaxed and amused, which improves their learning and understanding of the new language.
5. It increases the discussions and communications of the students with each other, they feel more relaxed and free in the class and it decreases the dominance of the teacher.

Limitations of the Study

The present study suffered from limitations beyond the researcher's control and, thus, may restrict the study's conclusions. Some other limitations may have occurred, which are collected as follows:

1. All learners do not access the internet and computer easily, and they have to use computers in public sites that would threaten the privacy of flipped classrooms.
2. The best performance in a flipped classroom depends on learners' motivation. Demotivated learners in these kinds of classes may perform lazily.
3. Learning must be done indefinitely because long-lasting learning in the flipped classroom may slow the learning speed and learners' interaction.

Delimitation of this Study

Some delimitations in this study threatened the generalizability of the results. The subjects of this research were Iranian students who were selected from Zanjan city. Thus, care should be taken to the generalizability of this study to other nationalities.

References

Ahadiat, E. (2019). Corruption and abuse of power: A reflection of social issues in short stories. *KnE Social Sciences*, 234-243. https://doi.org/10.18502/kss.v3i14.4311

Al-Seghayer, K. (2015). Salient key features of actual English instructional practices in Saudi Arabia. *English language Teaching, 8*(6), 89-99. http://dx.doi.org/10.5539/elt.v8n6p89

Amiryousefi, M. (2019). The incorporation of flipped learning into conventional classes to enhance EFL learners' L2 speaking, L2 listening, and engagement. *Innovation in Language Learning and Teaching, 13*(2), 147-161. https://doi.org/10.1080/17501229.2017.1394307

Andujar, A., Salaberri-Ramiro, M. S., & Cruz Martínez, M. S. (2020). Integrating flipped foreign language learning through mobile devices: Technology acceptance and flipped learning experience. *Sustainability*, *12*(3), 1110-1118. https://doi.org/10.3390/su12031110

Awidi, I. T., & Paynter, M. (2019). The impact of a flipped-classroom approach on the student learning experience. *Computers & Education*, *128*, 269-283. https://doi.org/10.1016/j.compedu.2018.09.013

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers & Education*, *78*, 227-236. https://doi.org/10.1016/j.compedu.2014.06.006

Bergmann, J., & Sams, A. (2012). Before you flip, consider this: Leaders of the flipped classroom movement say each teacher will have a different experience, but securing school leadership support, time, and IT resources will be important to every effort. *Phi Delta Kappan, 94*(2), 25-47. https://doi.org/10.1177/003172171209400206

Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Computers & Education*, *151*, 1-53. https://doi.org/10.1016/j.compedu.2020.103819

Boss, S., & Larmer, J. (2013). *PBL for 21st-century success: Teaching critical thinking, collaboration, communication, and creativity*. Buck Institution for Education.

Chan, S. Y., Lam, Y. K., & Ng, T. F. (2020). Student's perception of the initial experience of the flipped classroom in pharmacy education: Are we ready? *Innovations in Education and Teaching International*, *57*(1), 62-73. https://doi.org/10.1080/14703297.2018.1541189

Chance, P. (1986). *Thinking in the classroom: A survey of programs*. Teachers College Press.

Chen, S., She, J., Kameda, H., & Ohno, S. (2014, September). Implementation and evaluation of flipped classroom in the Chinese language course. In *Proceedings of Multidisciplinary Academic Conference* (pp. 1-8). https://doi.org/10.18026/cbayarsos.600312

Chen, Y., Wang, Y., & Chen, N. S. (2014). Is FLIP enough? Or should we use the FLIPPED model instead? *Computers & Education*, *79*, 16-27. https://doi.org/10.1016/j.compedu.2014.07.004

Coates, H. (2007). A model of online and general campus-based students. *Assessment & Evaluation in Higher Education, 32*(2), 121-141. https://doi.org/10.1080/02602930600801878

Cooper, K. S. (2014). Eliciting engagement in the high school classroom: A mixed-methods examination of teaching practices. *American Educational Research Journal, 51*(2), 363- 402. https://doi.org/10.3102/0002831213507973

Crouch, C., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970-977. https://doi.org/10.1119/1.1374249

Dehghanzadeh, S., & Jafaraghaee, F. (2018). Comparing the effects of traditional lecture and

flipped classroom on nursing students' critical thinking disposition: A quasi-experimental study. *Nurse education today*, *71*, 151-156. https://doi.org/10.1016/j.nedt.2018.09.027

Denprapat, O., & Chuaychoowong, M. (2016). Using flipped classroom model to develop English competency and independent attributes of Mathayom Suksa 1 students at Mengrai Maharajwitthayakhom school. *Journal of Social Academic, 9*(2), 225-239. https://www.tcithaijo. org/index.php/social_crru/article/view/80360/64034.

Dooly, M., & Sadler, R. (2020). "If you don't improve, what's the point?" Investigating the impact of a "flipped" online exchange in teacher education. *ReCALL*, *32*(1), 4-24. https://doi.org/10.1017/S0958344019000107

Dortaj, F (2004). *Investigating the effect of mental process simulation on improving students' academic performance and validating performance tests* [Unpublished doctoral dissertation]. Faculty of Educational Sciences and Psychology, Allameh Tabatabaei University, Tehran, Iran.

Elmaadaway, M. A. N. (2018). The effects of a flipped-classroom approach on class engagement and skill performance in a blackboard course. *British Journal of Educational Technology*, *49*(3), 479-491. https://doi.org/10.1111/bjet.12553

Evseeva, A., & Solozhenko, A. (2015). Use of flipped classroom technology in language learning. *Procedia-Social and Behavioral Sciences*, *206*, 205-209. https://doi.org/10.1016/j.sbspro.2015.10.006

Farah, M. (2014). *The impact of using flipped classroom instruction on the writing performance of twelfth-grade female Emirati students in the Applied Technology high school* [Unpublished doctoral dissertation]. British University in Dubai. http://bspace.buid.ac.ae/bitstream/1234/676/1/120088.pdf

Hamdani, M. (2019). Effectiveness of flipped classroom (FC) method on the development of English language learning of the high school students in Ahwaz. *International Journal of Applied Linguistics and English Literature, 8*(2), 12-20. http://dx.doi.org/10.7575/aiac.ijalel.v.8n.2p.12

Han, J., Huh, S. Y., Cho, Y. H., Park, S., Choi, J., Suh, B., & Rhee, W. (2020). Utilizing online learning data to design face-to-face activities in a flipped classroom: A case study of heterogeneous group formation. *Educational Technology Research and Development*, *6*,1-17. https://doi.org/10.1007/s11423-020-09743-y

Han, Y. J. (2015). Successfully flipping the ESL classroom for learner autonomy. *NYS TESOL Journal*, *2*(98), 98-109.

Henderson, K., Hobbs, C., & Gwynllyw, R. (2018, January). *Use of technology-enhanced teaching rooms to support flipped teaching* [Paper presentation]. The 12th International Symposium on Advances in Technology Education Nurturing Professionals for Smart Cities. https://isate2018.vtc.edu.hk/files/Full%20Paper/Parallel%20Session%201.pdf

Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning, 28*(1), 81-96. https://doi.org/10.1080/09588221.2014.967701

Johnson, L., & Renner, J. (2012). *Effect of the flipped classroom model on secondary computer applications course: Student and teacher perceptions, questions, and student achievement* [Unpublished doctoral dissertation]. Midwestern State University, Texas, USA.

Judy Shih, H. C., & Huang, S. H. C. (2020). EFL learners' metacognitive development in flipped learning: A comparative study. *Interactive Learning Environments*, *2*, 1-13. https://doi.org/10.1080/10494820.2020.1728343

Jungić, V., Kaur, H., Mulholland, J., & Xin, C. (2015). On flipping the classroom in large-first-year calculus courses. *International Journal of Mathematical Education in Science and Technology*, *46*(4), 508-520.

https://doi.org/10.1080/0020739X.2014.990529

Jungst, S., Wiersema, J., & Licklider, B. (2003). Providing support for faculty who wish to shift to a learning-centered paradigm in their higher education classrooms. *Journal of the Scholarship of Teaching and Learning, 3*(3), 69-81. https://scholarworks.iu.edu/journals/index.php/josotl/article/view/1606

Kang, N. (2015). The comparison between regular and flipped classrooms for EFL Korean adult learners. *Multimedia-Assisted Language Learning, 18*(3), 41-72. https://doi.org/10.15702/mall.2015.18.3.41

Kong, S. C. (2015). An experience of a three-year study on the development of critical thinking skills in flipped secondary classrooms with pedagogical and technological support. *Computers & Education*, *89*, 16-31. https://doi.org/10.1016/j.compedu.2015.08.017

Lai, C. L., & Hwang, G. J. (2016). A self-regulated flipped-classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, *100*, 126-140. https://doi.org/10.1016/j.compedu.2016.05.006

Liu, D. (2016). The reform and innovation of English course: A coherent whole of MOOC, flipped classroom, and ESP. *Social and Behavioral Science, 232*, 280-286. https://doi.org/10.1016/j.sbspro.2016.10.021

Lombardini, C., Lakkala, M., & Muukkonen, H. (2018). The impact of the flipped classroom in principles of microeconomics course: Evidence from a quasi-experiment with two flipped classroom designs. *International Review of Economics Education*, *29*, 14-28. https://doi.org/10.1016/j.iree.2018.01.003

Luo, H., Yang, T., Xue, J., & Zuo, M. (2019). Impact of student agency on learning performance and learning experience in a flipped classroom. *British Journal of Educational Technology*, *50*(2), 819-831. https://doi.org/10.1111/bjet.12604

Mason, G., Shuman, T., & Cook, K. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education, 56*(4), 430-435. https://doi.org/10.1109/TE.2013.2249066

McLaughlin, J. E., & Rhoney, D. H. (2015). Comparison of an interactive e-learning preparatory tool and a conventional downloadable handout used within a flipped neurologic pharmacotherapy lecture. *Currents in pharmacy teaching and learning, 7*(1), 12-19.

Murphree, D. S. (2014). "Writing wasn't really stressed, accurate historical analysis was stressed": Student perceptions of in-class writing in the inverted, general education, university history survey course. *History Teacher, 47*(2), 209-219.

Nicolosi, V. (2014, November 27-28). Analysis of aircrafts kinematic characteristics for pavement design. In O. Čokorilo (Ed.), *International Conference on Traffic and Transport Engineering (ICTTE)-Belgrade* (pp. 458-466). City Net Scientific Research Center.

Özkurkudis, M. J. M., & Bümen, N. T. (2019). Flipping the writing classroom: Using grammar videos to enhance writing. *Journal of Education and Future*, *15*, 1-16. https://doi.org/10.30786/jef.425632

Pakmehr, H., Mirdoraghi, F., Ghanaei, C. A., & Karami, M. (2013). Reliability, validity, and factor analysis of critical thinking disposition scales in high school. *Training Measurement, 4*(11), 33-53.

Rahman, M. M., & Alhaisoni, E. (2013). Teaching English in Saudi Arabia: Prospects and challenges. *Academic Research International*, *4*(1), 112-131. https://www.researchgate.net/publication/337768264_Teaching_of_English_in_Saudi_Arabia_Prospects_and_Challenges

Roach, T. (2014). Student perceptions toward flipped learning: New methods to increase

interaction and active learning in economics. *International review of economics education*, *17*, 74-84. https://doi.org/10.1016/j.iree.2014.08.003

Segolsson, M., Hirsh, A., & Bäcklund, J. (2017). The flipped classroom and student learning at compulsory school in Sweden: A longitudinal, qualitative study. *Journal of Education and Practice, 8*(18), 77-86.

Shahamat, H., Nadimi, H., Gharehbaglou, M., & Keramati, E. (2019). An analysis of the motivating factors underlying the hidden curriculum components through the lens of the Iranian students of architecture: A qualitative study. *Teaching and Teacher Education, 86*, 102-129. https://doi.org/10.1016/j.tate.2019.05.007

Sophia Learning, LLC. (2013). *Flip certified. Flipped classroom turns everything around*. http://www.sophia.org/flipped-classroom

Strayer, J. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research, 15*(2), 171-193. https://doi.org/10.1007/s10984-012-9108-4

Taylor, V., Tilhou, R., & Crompton, H. (2020). Computational thinking and coding across content areas to develop digital skills. In S. Yu, M. Ally, & A. Tsinakos (Eds.), *Emerging Technologies and Pedagogies in the Curriculum* (pp. 231-243). Springer.

Uzunboylu, H., & Karagozlu, D. (2015). Flipped classroom: A review of recent literature. *World Journal on Educational Technology*, *7*(2), 142-147. https://doi.org/10.18844/wjet.v7i2.46

Van Alten, D. C., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review, 28*, 1-8. https://doi.org/10.1016/j.edurev.2019.05.003

Walshaw, M., & Openshaw, R. (2011). Mathematics curriculum change: Parliament discussion over the past two decades. *Curriculum Matters, 7*, 1-6. http://www.questia.com/library/journal/1G1274585828/mathematicscurriculumchange-

Wang, T. (2020, January). Technology-enhanced multimodality in encouraging student interaction: A paradigm of a flipped second language classroom. *Advances in Economics, Business and Management Research, 116,* 293-298. https://doi.org/10.2991/icesed-19.2020.54

Wilson, S. G. (2013). The flipped class: A method to address the challenges of an undergraduate statistics course. *The Teaching of Psychology, 40,* 193-199. https://doi.org/10.1177/0098628313487461

Winch, J. (2020). An investigation of students' preferences in Japanese teaching and learning. *Global Journal of Foreign Language Teaching*, *10*(1), 72-84. https://doi.org/10.18844/gjflt.v10i1.4571

Yazici, H. J., Zidek, L. A., & Hill, H. S. (2020). A study of critical thinking and cross-disciplinary teamwork in engineering education. In A. E. Smith (Ed.), *Women in industrial and systems engineering* (pp. 185-196). Springer.

Zhonggen, Y., & Guifang, W. (2016). Academic achievements and satisfaction of the flipped business English writing class. *Educational Technology & Society, 3*(2), 298-312. https://eric.ed.gov/?id=EJ1097269

Zou, D. (2020). Gamified flipped EFL classroom for primary education: Student and teacher perceptions. *Journal of Computers in Education*, *7*(2), 213-228. https://doi.org/10.1007/s40692-020-00153-w

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

The Impact of Computerized Group Dynamic Assessment on Iranian EFL Learners' Listening Comprehension Across Gender

Research Article
pp. 205-224

Shahin Abassy Delvand[1]
Davood Mashhadi Heidar*[2]

Received: 2021/02/03 Accepted: 2021/04/20

Abstract
Framed in Vygotskian sociocultural theory, this study intends to examine whether computerized group dynamic assessment (GDA) through software has affected Iranian male and female learners' listening comprehension ability. Data were collected through administration of listening comprehension pre- and post-tests among 140 participants divided into male and female learners in the experimental and control groups. There were 35 male and 35 female learners in each group of the study. Participants in the experimental groups were exposed to GDA in order to interactively work on the selected tasks of listening comprehension, and the teacher provided the necessary support as well. Quantitative analysis of the pre- and post-tests of listening comprehension among male and female groups was conducted through two-way analysis of variance and covariance. Results revealed that both male and female learners in the experimental groups significantly outperformed the learners in the control groups. However, there were not any significant differences between the gender groups' listening comprehension ability in the experimental groups. Findings contributed to the effective employment of GDA through software in order to improve the learners' listening comprehension ability, denoting that teachers are recommended to be aware of technological devices in paving an interactive way for learners to develop their language skills and sub-skills.

keywords*:* computer assisted language learning (CALL), group dynamic assessment, EFL, listening comprehension ability, zone of proximal development (ZPD)

*Corresponding Author
[1] Ph.D. Candidate, Department of English, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran. Shahin.abacy@gmail.com
[2] Assistant Professor, Department of English, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran. d.mashhadiheydar@toniau.ac.ir

DOI: 10.22051/LGHOR.2021.34839.1440

Introduction

Over the last decades, much attention has been devoted to the learning, teaching, and testing of second language listening skill. This growing attention is due to the identification of the significance of listening in language learning. In spite of the significance of listening comprehension, English language classes, particularly in the public sectors, such as high schools of Iran, still focus on other skills more than listening ability. In addition, private language Institutes in Iran do not pay a well-deserved attention to teaching listening (Sehati & Khodabandehlou, 2017). In the Iranian EFL context, students "are not accustomed to hearing the language as it is produced by native speakers for native speakers" (Sehati & Khodabandehlou, 2017, p. 118). Accordingly, they have difficulties comprehending the language they hear when they are in contact with English speakers. In this regard, DA, which originated from Vygotsky's Socio-Cultural Theory (SCT), can be applied as both an evaluative and an instructional tool, offering new perspectives for assessment and teaching of listening comprehension. As Lantolf and Poehner (2004) state, Dynamic assessment integrates assessment and instruction into a seamless, unified activity aimed at promoting learner development through appropriate forms of mediation that are sensitive to the individual's (or in some cases a group's) current abilities. In essence, DA is a procedure for simultaneously assessing and promoting development that takes account of the individual's (or group's) zone of proximal development. (p. 50)

This view contrasts sharply with traditional testing practices that distinguished assessment and instruction and focus on assessment in a standardized and conventional manner (Haywood & Lidz, 2007). Some experts like Douglas (2010) consider DA as a form of alternative assessment; however, they are different as the main principle of DA is constructing the unknown from the known through more knowledgeable person mediation. Thus, some reported DA studies (e.g., Poehner & Lantolf, 2010; Rahimi et al., 2015) have considered private mediation sessions during which a student and a teacher have dialogical cooperation in order to overcome a problematic linguistic issue that cannot be solved by the student alone. This has resulted in the misinterpretation that DA is not appropriate for whole classes. As Anton (2009) states, "DA procedures are ideally administered individually, which makes this type of assessment time-consuming" (p. 579). As a result, many school teachers viewed DA as an unworkable concept (Guk & Kellog, 2007). To deal with such a problem, Poehner (2009) introduced G-DA, in which the mediator considers an individual ZPD as part of the whole group ZPD to progress all of them through interaction. Therefore, classroom activities focus on group ZPD, not an individual. The group ZPD is pointed out implicitly by Vygotsky himself as the "optimum time for teaching both the group and individual ZPD" (1998, p. 204). G-DA aims to address the group's ZPD and involve students in activities that can only be performed by pooling the learners' resources together in dialogical interaction.

With the emergence of educational technology and the subsequent progresses in corpus linguistics, particularly the corpora designed for language assessment and testing objectives (Park, 2014), a major turn can be witnessed in the assessment field. As to the significance of benefiting from incorporating technology in foreign language learning, teaching, and testing, teacher can benefit from technological devices in the learning environment. The influence of computer-assisted language learning (CALL) on listening skills development has been well investigated on the basis of the examination of teaching different second languages, such as Russian

language (Lebedeva et al., 2016) or English language (Abdolrezapour, 2019; Elfi, 2019; Khoshsima & Mozakka, 2017; Nachoua, 2012; Vahdat & Eidipour, 2016), among many others, confirming the constructive effects of digital technology on the learners' listening skills. However, the present study aimed to use digital technology in both teaching and assessment procedure, and examine its possible effects on the male and female EFL students' listening comprehension. To achieve this goal, three research questions have been proposed:

RQ1: Does computerized group dynamic assessment via CoolSpeech software have any statistically significant effect on the listening comprehension ability of male EFL learners at upper-intermediate level?

RQ2: Does computerized group dynamic assessment via CoolSpeech software have any statistically significant effect on the listening comprehension ability of female EFL learners at upper-intermediate level?

RQ3: Is there any significant difference in the listening comprehension ability of male and female EFL learners at upper-intermediate level affected by computerized group dynamic assessment via CoolSpeech software?

According to the above-mentioned research questions, the following hypotheses can be addressed as follows:

$H_0$1: Computerized group dynamic assessment via CoolSpeech software does not have any statistically significant effect on the listening comprehension ability of male EFL learners at upper-intermediate level.

$H_0$2: Computerized group dynamic assessment via CoolSpeech software does not have any statistically significant effect on the listening comprehension ability of female EFL learners at upper-intermediate level.

$H_0$3: There is not any significant difference in the listening comprehension ability of male and female EFL learners at upper-intermediate level affected by computerized group dynamic assessment via CoolSpeech software.

Review of the Related Literature

DA was developed on the basis of the SCT of language learning. There has been growing attention to DA among educators and psychologists in the last few decades. The foundation of social constructivist theory is based on the Zone of Proximal Development (ZPD). By zone, Vygotsky means "an area of exploration for which a learner is cognitively prepared, but requires assistance and social interaction" to develop (Borchelt, 2007, p. 2). Concerning the role of constructivism in language learning, it is assumed that learners can benefit from interactive language classroom in which they can be engaged in more creative activities focusing on the construction of knowledge. Vygotsky thinks that social learning precedes cognitive development. During this process, social interaction plays a crucial role. The more knowledgeable other refers to anyone who has a higher ability level than the learner whether a teacher or an adult (Jones & Brader-Araje, 2002). In fact, Vygotsky (1978) defined sociocultural theory as the influence of caregivers, peers, as well as society on the higher-order development of individuals. The advocators of SCT consider that learning is a mediated activity in which different mediational means or tools help learners enhance their thinking and knowledge (Shepard, 2001). As learning takes place interactively, the learning environment is as flexible as possible for students to do their tasks cooperatively. To this end, James (2008) stated that "knowledge is created and shared in expansive learning cycles" (p. 30), by providing adequate feedback within the students' ZPD.

The concept of scaffolding is also demanding when educator faces students' need to be supported by a more knowledgeable peer. In this regard, the teachers try to direct the learners' attention toward the target feedback and gradually help learners to do the correction themselves, which, as stated by Vygotsky, can be the sign of "transition from other-to-self-regulation" in which the learners are responsible for their learning materials (as cited in McKeough & Lupart, 2013, p. 152). The shift of responsibility from instructors to students is generally referred to as learner autonomy. As Godwin-Jones (2011) argues, technology implementation facilitates learners' autonomy through providing a great deal of materials for self-learning.

Following Vygotsky, second language inquiry considers DA as a procedure that integrates instruction and evaluation (Lantolf & Poehner, 2008). Lantolf and Poehner believe that "assessment and instruction are inseparable components of the same dialectical activity" and that "assessment and instruction become as tightly conjoined as two sides of the same coin, and there are no one-sided coins" (p. 274). DA involves Interactionist and Interventionist types that commonly comprise three phases: pretest → mediation → posttest (Lantolf & Poehner, 2004). The interventionist kind of DA consists of the examiner's intervention throughout the procedure of assessment and it is a more standardized and formal approach. In this type of assessment, the examinees are instructed item by item and if they are not able to solve the items properly, they receive planned hints. The Interactionist DA, however, includes mediation, which emerges from the interaction between the examinee and examiner. In this type of assessment, hints, prompts, or leading questions are not planned beforehand; rather, they emerge from collaborative interaction or mediated dialogue between the examinee and the examiner where the examiner responds to the examinee's requirements. In DA, the relationship between examinee and examiner is based on the notion of helping and teaching, for example, students are permitted to ask questions and obtain feedback immediately. In both types of DA, teaching can be delivered in group or individual contexts (Poehner, 2005).

Prior to accounting for DA and G-DA, the role of gender in language instruction is worth investigation. There has been a common post-modern belief that gender is built on the foundation of social communications, and the gender difference was produced by the social arrangements and what a human being acquire by repeated actions in society. It is also found in conversation analysis (CA) that social behavior is determined by including all structures and interactions meaningfully (Cameron, 2005). In different contexts with different cultures, there are different beliefs of gender groups and their influence on language learning. The sociolinguistics message is that gender can influence the linguistic performance (Baxter, 2003).

The language and gender relationships in educational context shows that in the past, the role of women in public positions was denied and even women did not have the right to access to language, but today, although some of that forbiddance is removed, but some serious power still remained. Based on different studies (Walsh, 2001) women performed well in different situations and their linguistic behavior reflects their understanding, because they believe that they should put more effort on showing their strict to norms. Also it is believed in countries with advanced economics that the main criterion for employing of workforce is now interpersonal and linguistic, not physical strength. Hence, the role of gender appears to have been taken for granted in issues related to language teaching and learning affected by

different interventions, such as DA and G-DA, which are the focus of the current study.

The support for implementing both types of this assessment (DA and G-DA) in language education, and particularly in the assessment of listening skill, has grown in previous years. For example, a research conducted by Ableeva (2008) applied the DA practice to listening testing to report the possible contributions of such assessment to the listening comprehension instruction and assessment. The study followed the pretest-mediation-retest format. For the pretest, some recording was played twice for the students and they were required to write the responses. During the intervention, the students could ask questions, and the mediator provided suggestions, explanations, hints, and other mediations. Throughout the retest phase, students were required to write a summary of the text. The researcher concluded that DA was a valuable diagnostic tool that empowered her to recognize the particular source of the difficulty that was not obvious during a conventional test. Furthermore, this evaluation could not only enable teachers to diagnose the actual level of students' listening comprehension ability but also to indicate their potential capabilities that were in the process of developing.

Hidri (2014) also conducted a study in which 60 Tunisian EFL learners answered to a listening comprehension assessment with 2 parts, dynamic and static. The dynamic part of the test administration was conducted during regular hours of progress assessment in the class; whereas, the learners' final achievement tests was fulfilled through the static assessment, occurring after one year of listening practice. DA consisted of 3 assessment phases which were administered in 45 minutes. The tests comprised 14 items that aimed to create negotiation of meaning between the mediators and test-takers. The mediators provided guidance and support in the pre- and while-testing stages. However, they were trained to decrease mediation during the posttest. The pretest involved matching, guessing, and WH- items. The testing phase included guessing, true/false, multiple choices, summarizing, and wh-questions. The posttest contained making inference, summarizing, picture reordering, and multiple-choice items. The achievement static test involved 40 true/false, information transfer, multiple-choice, and gap-filling items. Eleven raters scored the tests. Both the raters and test-takers were asked about their attitudes towards the two assessments. The analyses demonstrated that item difficulty estimates, rater behavior, and test-taker ability differed across assessment types. The findings demonstrated that "although the new assessment provided better insights into learners' cognitive and meta-cognitive processes than did the traditional assessment, raters were doubtful about the value of and processes involved in DA mainly because they were unfamiliar with it" (p. 1).

In a similar vein, Mashhadi Heidar and Afghari (2015) explored the listening skills of Iranian EFL students at the upper-intermediate level through examining the effects of DA in a computer-mediated environment. These researchers investigated 60 students' socio-cognitive development through DA. The findings demonstrated that through interaction in the ZPD, DA in synchronous computer-mediated communication provided the opportunity to explore the actual level of the students' listening ability, as well as to identify and evaluate the potential level of their development.

In a different yet relevant study, Mashhadi Heidar (2016) examined the role of DA in improving the listening skills of 30 EFL students at the intermediate level via web 2.0. The control group experienced conventional techniques, whereas the

experimental group was exposed to technology-based intervention mediated instruction. Firstly, the learners performed the recall without mediation (NDA phase) and afterwards they repeated it, with mediation (DA phase). Thus, two opportunities existed to evaluate the listening improvement during the DA and NDA sessions, including mediated and independent listening performance. The findings of the study demonstrated that technology-based DA improved the participants' listening comprehension skill. It was also indicated that through online DA, the source of poor performance can be identified, which is typically hidden during non-dynamic ones.

Regarding the possible impacts of computerized DA on the listening skill development of students, Ashraf et al. (2016) explored the effect of computerized DA on the listening ability of Iranian EFL students. To this end, they administered a Quick Placement Test to 65 EFL learners in a language institute, and selected 40 upper-intermediate students as the participants of their study. The selected students were administered a listening test as the pretest and posttest. They were then divided into a control and an experimental group. The experimental group received the instruction via electronic-based DA in a virtual class, while the control group was exposed to the listening instruction via traditional DA in a physical classroom. The results demonstrated that the electronic-based DA could significantly affect the learners' listening comprehension.

Another similar study that investigated G-DA of EFL learners' listening ability was conducted by Ahmadi Safa and Beheshti (2018), who investigated the effect of interventionist and interactionist approaches to GDA on listening comprehension development of Iranian EFL students. Ninety intermediate learners were assigned to two experimental groups and one control group, which were then divided into subgroups. The interactionist approach was applied in the first experimental group, where the teacher took part in the listening comprehension activities, interacting, and assisting the group members in their activities. The interventionist approach to DA was used in the second experimental group, where the teacher participated in the activities and provided the students with feedback. The control group, however, used the conventional summative assessment. The data analyses indicated the primacy of interactionist GDA for enhancing the students' listening skills. Furthermore, although the interventionist GDA procedure had more effectiveness than the non-dynamic procedure of the control group, the difference was not significant. It was concluded that interactionist GDA was an effective practice and that the authoritative and unilateral approaches to pedagogy need to be replaced by the cooperative and interactive ones.

The effect of the G-DA approach on the listening skill of Iranian EFL learners was investigated by Roohani et al. (2018) in two time intervals. They selected 20 learners, who answered a pretest individually. After collecting the test sheets, the researchers replayed the listening tests for the students to provide their recalls in order to perform the test in a dynamical format. The listening tasks were provided in 3 sessions. During each session, the students first performed a listening task non-dynamically, then, they performed that listening task through G-DA practice. After this phase, the participants took the posttest to reveal their listening comprehension development. The researchers' data analysis indicated an enhancement in the participants' posttest scores, and they concluded that the improved listening comprehension could primarily be attributed to the mediation provided throughout G-DA.

Alshenqeeti and Grami (2019) also examined the effect of DA in listening comprehension classrooms. A total of 56 Arab EFL learners participated in their study. The participants took a general placement proficiency test to confirm their homogeneity. The pretest and posttest scores of the students were recorded and the values were coded quantitatively. One-way ANOVA was used to analyze the outcomes. The researchers supported the effect of DA in improving listening ability through the fact that DA participants' scores were better than their counterparts. The researchers also suggested language teachers to use mediation practices to enhance the learners' listening abilities.

Generally, the literature review related to the assessment of listening skills reveals that various investigations have been conducted to address different assessment approaches. However, most of the investigations on L2 listening have concentrated on improving listening comprehension, methodologies to teach listening, the appropriateness of listening materials, and similar matters. Few have focused their attention on the assessment of listening comprehension in general and computerized G-DA method to evaluate this skill in particular. Besides, as Alderson and Bachman (2003) argue, assessing listening skills is one of the least developed, least understood, and yet one of the most important areas of language assessment and testing. Actually, the review of the literature yielded that the present study is one of the few attempts that specifically compare female and male learners in computerized group dynamic approaches to assessment in the area of listening comprehension, particularly in the Iranian EFL setting. Further, sometimes in different contexts, the findings might be rather contradictory. For instance, several researchers (Ableeva, 2008; Alshenqeeti & Grami, 2019; Ashraf et al., 2016; Hidri, 2014; Roohani et al., 2018) suggested that the implementation of DA could support the students' listening comprehension achievement. Ghahremani (2013) also compared 3 types of assessment (Dynamic, formative, and summative assessment) on listening skill and listening strategy use of 140 Iranian university students and concluded that the participants in the dynamic group showed a better performance than the other two groups. It was suggested that summative assessment is important but it does not support learning and provides no continuous and planned feedback during the learning process. However, other researchers (e.g. Harlen, 2006; Lam & Lee, 2010; Lee & Coniam, 2013) agree that when using different approaches to assessment, the focus is greatly on summative assessment rather than doing the assessment in the learning process by involving the learners. Actually, summative assessment has a long history in education, and parents and the public have widely accepted them. In fact, it is regarded as an indispensable part of the language learning process in academic activities for scoring, record-keeping, and reporting (Brookhart, 2008). Moreover, due to the exam-oriented culture in the Asian academic system, alternative forms of assessments might not be able to substitute summative assessments (Kennedy et al., 2008). Obviously, as researchers make their way into this field of inquiry, challenges continue to emerge. Therefore, the current investigation tried to address the gap in research studies and aimed to scrutinize the effect of computerized G-DA on male and female EFL learners' listening comprehension ability.

Methodology

Participants

To collect the data, 140 Iranian EFL language learners, male and female, from Kish Language Institute were considered as the major participants of the investigation. They were selected through convenience sampling from among adult students of the Institute because the researcher aimed at keeping age as a fixed variable, then they were assigned randomly to the experimental and control group. Teenager participants needed different approaches of treatment. The same problem existed for middle-aged participants. That was why the age range of the participants was between 22 and 35. To make sure of homogeneity, participants were chosen from among 200 students, based on their results on an Oxford Placement Test (OPT). One hundred and forty upper-intermediate students were selected from among 200 potential participants using OPT. Then, the first experimental group included 35 male students; the second experimental group similarly involved 35 female students. In addition, the first and second control groups of the study included 35 male and 35 female participants, respectively. Table 1 shows the participants' demographic information.

Table 1
Demographic Background of the Participants

No. of Students	140 upper inter-mediate learners (35 experimental one; 35 experimental two; 35 control one; 35 control two)
Gender	70 Females & 70 Males
Native Language	Persian
Age Range	22-35 years old
Institute	Kish Language Institute
Academic Years	2019-2020

Instruments and Materials

The following instruments were applied for collecting the required data.

Oxford Placement Test (OPT). To ensure the participants' homogeneity, they were chosen based on their results in an Oxford Placement Test (OPT). As OPT is a standardized exam, the standardization process was not necessary. The OPT which was used in the current research consisted of 60 multiple-choice items. Participants had 30 minutes to answer the questions. The students who gained scores from 40 to 47 were considered as upper-intermediate level or B2.

Listening Comprehension Test. In order to examine the learners' listening comprehension ability, a 20-item researcher-made listening comprehension test was used. It was designed according to the textbook taught at Kish Language Institute. It consisted of 5 listening tasks each followed by 4 multiple comprehension choice questions. It was designed for upper-intermediate level of proficiency. In order to make sure that the listening items in the tests were of the right level, they were selected from different listening tasks of the textbook covered at the institute. The tests were made reliable by pilot study explained as follows.

In the phase of pilot study, the researcher designed these 3 listening tests and gave each of them to 15 participants, who were the representative of the participants of the study. The reliability coefficient was calculated as .71 through Cronbach's alpha. The same listening comprehension test was used as both the pre- and post-test in order to avoid the risk of different levels of difficulty in different

versions of the test.

CoolSpeech Software (latest version 5.0.) CoolSpeech Software was developed by ByteCool Software Inc. (2001). This text-to-speech program is supported by the most current Windows systems. This software can be applied to convert any sorts of text files (.txt, .rtf, and .htm/Html) into voice files (.Wav), and read the texts (from a variety of sources) aloud. The following features can be specified for this software:

- Being free of charge and user-friendly;
- Online listening to the news provided from any URLs;
- Reading texts of any common formats aloud;
- Fostering independent learning;
- Listening to every expression the user has just typed anywhere in Windows; and
- Audio-file scheduling, which can be listened further as a practice

Design

The current investigation was a quasi-experimental research in which the participants' homogeneity was initially ensured through administering OPT, and afterwards, they were assigned into two experimental and two control groups to meet the research purpose. The study employed a pretest, treatment, and posttest design.

Data Collection Procedure

After OPT administration and ensuring the participants' homogeneity in terms of proficiency level, then, they were assigned into 2 experimental and 2 control groups as male upper-intermediate learners, and female upper-intermediate learners. As the researcher wanted to have the groups of equal numbers, 35 male and 35 female learners were selected and assigned into two experimental groups. Moreover, 35 male and 35 female participants were assigned into 2 control groups, and the rest were excluded.

Prior to the treatment sessions, the homogeneity of participants' listening comprehension was checked by administering the pretest. Pretest results revealed no significant differences among the study groups, ensuring listening comprehension homogeneity among the participants of the study.

Both experimental groups underwent the same treatment of using computerized G-DA through CoolSpeech software for 30-minute 30 sessions. The target listening comprehension tasks were selected based on the learners' level of proficiency, encouraging them to work in pairs and groups. CoolSpeech software was used in order to develop their listening comprehension ability with the same tasks, and the learners benefited from group discussions on each task. Monitoring the learners' interactions was conducted by the teacher who provided necessary feedback as well. However, an attempt was made to promote self and peer correction rather than teacher feedback. In other words, the teacher acted as the facilitator who paved the way for learners to improve their listening comprehension in an interactive learning environment.

In the experimental groups, the following mediational tactics (Lantolf & Poehner, 2008) were used. Prior to the instruction, the teacher provided specific explanations regarding CoolSpeech software and how the learners are to work with this software during the treatment sessions. It is worth mentioning that this software

was used as a supplementary tool in line with G-DA to pave the way for EFL learners to develop their listening comprehension. Through this software, learners were able to work independently on each listening task and the teacher could save the time more efficiently and allocate more effective time for each student to carry out the provide tasks more actively. CoolSpeech was employed in order to facilitate G-DA in the sense that: First, the teacher confirmed the correct responses that the students were not certain about. Second, whenever it was needed the instructor allowed the students to hear the passage again through CoolSpeech software; participants were armed with a computer on which CoolSpeech software was installed. Third, whenever students could not comprehend the text after replaying it, the instructor tried to divide sentences into smaller and more intelligible chunks, in order to make them more comprehensible to the students. Fourth, the instructor restated the students' wrong statements with a questioning tone, acting as a model to show learners how to do the corrections and be the assessor. Fifth, learners were requested to provide peer feedback and be more engaged in doing the listening tasks more communicatively and provide correction for their classmates. Sixth, the instructor offered learners contextual clues, including topical knowledge, world knowledge, and situational awareness. Seventh, the instructor mediated students through using metalinguistic clues, either grammatical or lexical clues. Eighths, the instructor allowed the students to apply a dictionary, whenever they could not guess a word from its context. Finally, whenever the other mediational tactics could not help the students to guess the correct response, the instructor explained it.

In contrast to the experimental groups, CoolSpeech software and G-DA were not employed in the control group, receiving the conventional method of teaching listening comprehension in which the teacher provided the listening comprehension exercises followed by a warm-up. In fact, no computerized G-DA was used for the control group. Then, the recording was played and the learners were encouraged to discuss the comprehension questions posed by the teacher. Any ambiguities in understanding the task were resolved by the teacher. The learners listened to the task for the second time in order to have more concentration and resolve any misunderstanding. The learners' grades were much focused in the control group.

After 3 months of treatment on the experimental groups, which took 30 sessions, 30 minutes each session, the participants in all groups took part in a listening posttest, and the results were compared and contrasted to check the hypotheses of the study.

Variables of the Study

As to the purpose of the study aiming to probe the effect of computerized G-DA through CoolSpeech software on Iranian EFL learners' male and female learners' listening comprehension ability, computerized G-DA, CoolSpeech software, and gender are considered as the independent variables of the study. However, listening comprehension ability involves the dependent variable of the study, which is affected by the above-mentioned independent variables.

Data Analysis

The current research employed quantitative methodology to answer the research questions of the study. Descriptive and inferential measures were used accordingly. The former is concerned with the learners' development of mean scores

from the pretest to the posttest between male and female groups. The latter includes one-way ANCOVA to investigate the possibility of significant difference between gender groups. Finally, the researchers employed two-way ANOVA to examine listening comprehension of the experimental group among male and female learners.

Results

Addressing First Research Question

The first purpose of the study was investigating upper-intermediate male learners' listening comprehension affected by computerized G-DA. In doing so, descriptive and inferential measures were conducted in order to compare the pretest and posttest scores of listening comprehension. Table 2 shows descriptive statistics for the listening comprehension ability of male participants in the experimental and control groups.

Table 2

Descriptive Statistics for the Male Learners' Listening Pre- and Post-Tests in the Experimental and Control Groups

	N	Mean	Std. Deviation	Std. Error Mean
Experimental post	35	70.60	3.776	1.192
Control post	35	65.71	5.458	1.726
Experimental pre	35	65.86	4.995	1.803
Control pre	35	64.90	5.109	1.119

Table 2 shows that there is a negligible difference in the pretest scores of the male experimental (M = 65.86; SD = 4.99) and control (M = 64.90; SD = 5.10) groups. However, as to the post-test results, Table 2 demonstrates the existence of a fairly large difference between the posttest mean scores of the experimental group (M = 70.60; SD = 3.77) and the control group (M = 65.71; SD = 5.45) for the male participants. In order to inferentially compare the experimental and control learners' listening comprehension ability after exposure to computerize G-DA, one-way analysis of covariance (one-way ANCOVA) was run as in Table 3.

Table 3

ANCOVA Results for Male Learners' Listening Posttest in the Experimental and Control Groups

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power
Corrected Model	433.517a	2	216.759	44.379	.000	.839	88.758	1.000
Intercept	29.709	1	29.709	6.083	.025	.264	6.083	.643
test writing-male-pre	313.467	1	313.467	64.178	.000	.791	64.179	1.000
Code	76.860	1	76.860	15.736	.001	.481	15.736	.962
Error	83.033	67	4.882					
Total	93405.000	70						
Corrected Total	516.550	69						

a. R Squared = .839 (Adjusted R Squared = .820)

b. Computed using alpha = .05

As shown in Table 3, the significance level is less than .05 ($F_{1, 32} = 15.73$, p = .001), showing a significant difference between the experimental group and the control one among the male participants. Therefore, it can be concluded that listening instruction through computerized G-DA could result in significant improvement in the Iranian EFL male learners' listening comprehension ability, thus leading to the rejection of the first null hypothesis.

Addressing Second Research Question

The second research question of the study aimed to examine the effect of computerized G-DA on female learners' listening. Table 4 provides descriptive information regarding the female participants' listening performance in the experimental and control groups pre- and post-tests.

Table 4

Descriptive Statistics for the Female Learners' Listening Pre- and Post-Tests in the Experimental and Control Groups

	N	Mean	Std. Deviation	Std. Error Mean
Experimental post	35	69.60	6.758	1.746
Control post	35	62.17	4.166	1.076
Experimental pre	35	63.20	5.212	1.179
Control pre	35	63.10	5.111	1.201

Descriptive statistics for the pre-test scores of the experimental (M = 63.20; SD = 5.21) and control groups (M = 63.10; SD = 5.11) in Table 4 reveals a small difference of female participants' listening. However, the Table shows almost much difference between the two group, since the experimental group (M = 69.60; SD = 6.75) performed better than the control group (M = 62.17; SD = 4.16). The examination of the significance level between the two groups was done through inferential one-way ANCOVA as in Table 5.

Table 5

ANCOVA Results for the Female Learners' Listening in the Experimental and Control Groups

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power
Corrected Model	1224.404a	2	612.202	197.335	.000	.936	394.672	1.000
Intercept	1.355	1	1.355	.438	.514	.016	.437	.098
test writing-female-pre	798.770	1	798.770	257.473	.000	.905	257.474	1.000
Code	190.731	1	190.731	61.480	.000	.695	61.480	1.000
Error	83.763	67	3.102					
Total	131329.000	70						
Corrected Total	1308.167	69						

a. R Squared = .936 (Adjusted R Squared = .931)

b. Computed using alpha = .05

Table 5 indicates that the significance level is less than .05, verifying a significant difference between the female learners' performance ($F_{1, 32} = 61.48$, p = .000). In other words, independent samples t-test revealed that listening instruction through computerized G-DA resulted in the female learners' significant

improvement in listening comprehension ability, resulting in the rejection of the second null hypothesis of the study.

Addressing Third Research Question

The third research question of the study aimed to take into account male and female learners' comparison of listening performance affected by the treatment. In doing so, Table 6 compares the descriptive statistics between the male and female participants who underwent computerized G-DA in listening instruction.

Table 6

Descriptive Statistics for the Male and Female Learners' Listening in the Experimental Group

	N	Mean	Std. Deviation	Std. Error Mean
Female	35	69.60	6.758	1.746
Male	35	70.60	3.776	1.192

Table 6 demonstrates the similar performance of the male (M = 70.60; SD = 3.77) and female (M = 69.60; SD = 6.75) groups on their listening posttest. The difference between the two groups' posttest was considered inferentially through a two-way analysis of variance (two-way ANOVA) in Table 7 below to compare the two gender groups' listening comprehension ability.

Table 7

Two-way ANOVA Results for the Male and Female Learners in the Experimental Group

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.000	1	6.000	.180	.676
Within Groups	768.000	68	33.391		
Total	774.000	69			

Table 7 presents that the significance level is more than .05 ($F_{3, 66} = .180$, p = .676), revealing no significant differences between the male and female groups' mean scores of the listening posttest. In other words, there is not any significant difference between the effect of listening comprehension through computerized G-DA regarding male and female learners' listening comprehension ability.

In sum, the results of the study acknowledged the upper-intermediate male and female learners' significant improvement in listening comprehension through computerized G-DA. However, there was no significant difference across gender groups.

Discussion

The current inquiry was an attempt to study the potential effects of the computerized G-DA method on the listening comprehension ability of male and female EFL students. To this end, four groups (two experimental and two control groups) formed the sample of the investigation. The results demonstrated the outperformance of the experimental groups compared to the control groups regarding their listening comprehension ability. Concerning the first research question, which investigated male EFL learners' listening comprehension affected by of G-DA via CoolSpeech software, it was found that computerized G-DA could help learners to gain mastery over listening comprehension tasks. With regards to

the second research question, which explored the female learners' listening as a result of applying computerized G-DA, results indicated that female learners' listening comprehension was also affected by such assessment. Concerning the final research question, which comparatively examined learners' listening comprehension ability affected by computerized G-DA in two gender groups, the results of the two-way ANOVA revealed no significant difference regarding their listening comprehension ability. Likewise, the learners in both experimental groups performed better than those of the control group. In fact, it appears that gender as an independent variable in this study was affected by the use of computerized G-DA as the main intervention in the sense that both male and female groups could significantly improve their listening comprehension ability.

The present study was framed in Vygotskian Socio-Cultural Theory. Vygotsky's SCT provides three main themes concerning the ZPD, the more knowledgeable other and social interaction. By zone, Vygotsky means an exploration area for which a learner has cognitive preparation, but needs assistance and social interchange to develop. ZPD concept can be incorporated in the language learning classroom. This needs the use of assistance and supporting resources to help learners, bridge the learning gap between the known and unknown course content. The scaffolding strategy is evident in the technology-based language learning classroom, because this instructional model requires the facilitator to provide learners with meta-cognitive support and ensure the exactness of student learning so that the learner can become independent. In Vygotsky's idea, knowledge is co-constructed through interaction and individuals learn from each other; correspondingly, a learner in this theory must be engaged in the learning process. All these can support the findings of the present inquiry, which revealed that the learners' listening comprehension was enhanced by G-DA. The students' improved scores on the posttest could be due to the mediation provided during the G-DA procedure and the use of CoolSpeech software. It has to be mentioned that this software could help the teacher to optimize the quality of G-DA by directing the learners' attention to the listening tasks more purposefully and encourage them to work independently. Furthermore, when an activity was misunderstood for one learner, he/she was able to replay the audio for himself/herself and then be engaged in classroom communications. Although G-DA can be practical on its own, applying the treatment with CoolSpeech could motivate learners' more participation in the classroom interaction and feel that they own the learning environment and be more autonomous in carrying out the listening tasks and enjoy online listening practice as an extra activity. Last but not least, the above-mentioned software could be beneficial for reticent learners who might feel less confident in taking part in classroom talk and gradually by working with the software and doing the listening tasks individually, they were able to increase their sense of peer work and participate in G-DA intervention.

The findings of this study are in alignment with those of Ableeva (2008), who applied the DA practice to listening comprehension assessment, and concluded that DA was an effective diagnostic instrument that enabled the teacher to identify the actual level of learners' listening comprehension ability and assist them in the process of developing their proficiency. The result related to the impacts of the G-DA on the listening skills of the students in the experimental group is also consistent with the inquiry by Roohani et al. (2018), who observed the effect of the G-DA approach on EFL learners' listening comprehension and indicated an enhancement in

the participants' posttest scores. The significance of DA in improving EFL learners' listening ability was also supported by Alshenqeeti and Grami (2019) through the fact that DA participants' results were better than their counterparts. The outcomes of the present investigation are also in line with the preceding studies on the positive influence of the DA in a computer-mediated environment, such as the investigations conducted by Mashhadi Heidar and Afghari (2015), and Mashhadi Heidar (2016), who explored the learners' socio-cognitive progress through electronic-based DA, and concluded that technology-based DA improved the learners' listening comprehension ability. Similarly, Ashraf et al. (2016) explored the effect of computerized DA on the listening skill of EFL students and demonstrated that this kind of assessment could significantly affect the learners' listening comprehension. The present study revealed that the cooperative and interactive approaches to pedagogy in G-DA led to the effectiveness of this method for both experimental groups. This is something which is in agreement with Ahmadi Safa and Beheshti's (2018) study, which investigated the effect of interventionist and interactionist approaches to G-DA on listening comprehension development of EFL students, and proved the primacy of interactionist G-DA for enhancing the students' listening skills, suggesting that the authoritative and unilateral approaches to pedagogy be replaced by the cooperative and interactive ones.

Despite the fact that gender has been regarded to play a momentous role in L2 acquisition, and educational research has confirmed the effect of gender differences on learners' academic achievements, needs, and interests, the findings of the present study demonstrated no substantial differences across gender groups in using G-DA procedure. It is worth mentioning that different academic areas have different declarations to the gender issues. As to the outcomes of the present investigation, it can be concluded that the G-DA procedure and its beneficial guidance might be used in both female and male classes in the same way. Moreover, the effective guidance of G-DA might be used in students' textbooks with no consideration of gender differences. More importantly, computerized G-DA can be successfully applied to both male and female groups as the findings of the study demonstrated that gender and development of listening comprehension were significantly correlated regardless of showing similar performance to the treatment. Therefore, gender has to be taken into consideration seriously an independent variable and in EFL instruction of language skills and both quantitative and qualitative researches are demanding to highlight the role gender plays in language teaching and learning.

Conclusion and Implication

As to the findings, it can be concluded that the G-DA method of instruction via CoolSpeech software could have a significant effect on the listening comprehension skill of both male and female learners, though no statistically significant difference between the two experimental groups (male experimental and female experimental group) was observed. The present study can have practical implications for the English language teaching and testing field.

The findings of the study can pedagogically contribute to the curriculum development and syllabus design regarding the accommodation of more interactive materials in order to foster further teacher-learner and learner-learner interactions in the classroom. It seems necessary to revise pedagogy to balance the tensions between dynamic and non-dynamic forms of assessment, and to use the advantages

of each to improve teaching and learning. One of the challenges to a learning-oriented approach to assessment is the views that simply equate evaluation with grading (Zeng et al., 2018). Dynamic approaches to the skills evaluation could help promote the features of assessment tasks as learning tasks and student involvement in the evaluation process. Exploring gender-based differences between students in applying G-DA practice can be valuable for policy makers. Moreover, the present outcomes seem to be beneficial for learners, teachers, teacher educators, and researchers.

Learners can positively be exposed to G-DA listening comprehension instruction through CoolSpeech software in order to develop their listening comprehension and enjoy more communication in the classroom, resulting in quality learning of listening comprehension. In comparison with the conventional type of assessment focusing on the overall listening ability of the learners, the dynamic model of assessment, as the present study suggests, can provide valuable insights for the learners to have awareness regarding their quality listening in the learning process. The learners' involvement in the assessment procedures not only motivates the learners' development of language skills, but they also help learners to be aware of the significance of being assessed during the learning process (Archer, 2010).

Teachers can also employ G-DA through CoolSpeech software in order to provide assessment-oriented instruction of listening comprehension for the learners in the communicative classroom. G-DA can be considered as a quality methodology to involve the learners in the listening comprehension tasks and encourage learners to have self and peer-corrections under the teacher's monitoring. The result of this study can lead both teachers and learners to apply the best assessment tools in the classroom in order to eliminate possible difficulties they may encounter during the listening class. In other words, dynamic assessment approaches can be adopted as a teaching methodology for teachers to trigger the learners' listening capacities and help them identify their weaknesses and strengths for the purpose of quality learning, and finally listening ability. Moreover, teachers need to provide assessment feedback to maximize its potential for students' actions in line with the DA framework.

The present study can also contribute to teacher educators' provision of in-service workshops for both novice and experienced teachers to benefit from assessment-oriented instruction in the classroom for the purpose of enhancing the teaching performance by involving the learners in the assessment procedure and help them to have more interaction with their peers by doing group DA.

Finally, the results of the study can be practically significant in terms of conducting further assessment-oriented studies, particularly in the Iranian context in which teaching listening comprehension through G-DA appeared to be an under-researched area of study.

Generalizing the findings of the study has to be carefully attended as the sampling and context setting do not represent the entire population of foreign language learner in Iran. The study was also limited to the context of a private language institute for participants' selection. Moreover, the participants constituted a small number of language learners. Another limitation was time, since the data collection occurred over the course of a few months. However, it is difficult to know whether this amount of time was enough to witness changes. Furthermore, a single, commonly agreed-upon definition of comprehension remains elusive (Cutting & Scarborough, 2006). Different comprehension assessments do not always generalize

across items, formats, and subjects due to differing definitions of comprehension. Thus, possibly the results of the current study might not generalize beyond the measures of listening comprehension used in this study. The final limitation of the study is related to the sampling procedure adopted in this study, since the researchers could not conduct a truly randomized sampling due to administrative reasons.

Thus, further research can be carried out to explore other variables, such as different learning environments, including high schools or academic settings like universities, as well as different levels of proficiency, and other language skills. In addition, further research can be done with the participation of a larger sample. Moreover, the teachers' perceptions about computerized G-DA in effective teaching of language skills can be explored. Investigating the teachers' perceptions can determine whether they are aware of practically applying G-DA or they should be provided with awareness-raising activities by the teacher education. It is also recommended to implement computerized G-DA in teaching other language skills and sub-skills; however, a pilot study is required to check the applicability of assessment types. Teachers and researchers can also consider an action research approach or mixed methods research that integrates philosophical, qualitative, quantitative, and action research for the study of different assessment scenarios. Last but not least, the learners' personality types (i.e., introversion/extroversion and reflectivity/impulsivity), their motivation, critical thinking, etc. can be taken into account in further research to examine how these groups of learners performed in assessment-based instruction. These factors might affect the students' performance in learning, and employing the DA method might produce thought-provoking results in terms of how to efficiently apply this approach in the classroom.

References

Abdolrezapour, P. (2019). Applying computer-mediated active learning intervention to improve L2 listening comprehension. *Applied Research on English Language*, *8*(4), 511-530.
http://dx.doi.org/10.22108/are.2019.115355.1424

Ableeva, R. (2008). The effects of dynamic assessment on L2 listening comprehension. In J. P. Lantolf & M. Poehner (Eds.), *Socio-cultural theory and the teaching of second languages* (pp. 57-86). Equinox Press.

Ahmadi Safa, M., & Beheshti, S. (2018). Interactionist and interventionist group dynamic assessment (GDA) and EFL learners' listening comprehension development. *Iranian Journal of Language Teaching Research*, *6*(3), 37-56.
https://doi.org/10.30466/IJLTR.2018.120600

Alderson, J. C., & Bachman, L. F. (2003). Series editors preface. In G. Buck, *Assessing listening* (pp. x-xi). Cambridge University Press.

Alshenqeeti, H., & Grami, G. M. A. (2019). Dynamic assessment in the EFL classroom: The case of listening comprehension. *Asian Journal of Education and Social Studies*, *5*(4) 1-11.
https://doi.org/10.9734/ajess/2019/v5i430160

Anton, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals*, *42*(3), 576-598.
http://dx.doi.org/10.1111/j.1944-9720.2009.01030.x

Archer, J. C. (2010). State of the science in health professional education: Effective feedback. *Medical Education, 44*, 101-108.
http://dx.doi.org/10.1111/j.1365-2923.2009.03546.x

Ashraf, H., Motallebzadeh, K., & Ghazizadeh, F. (2016). The impact of electronic-based dynamic assessment on the listening skill of Iranian EFL learners. *International Journal of Language Testing, 6*(1), 24-32.
https://www.ijlt.ir/article_114420.html

Baxter, J. (2003). *Positioning gender in discourse*. Palgrave.

Borchelt, N. (2007). Cognitive computer tools in the teaching and learning of undergraduate calculus. *International Journal for the Scholarship of Teaching and Learning, 1*(2), 1-17.
https://doi.org/10.20429/ijsotl.2007.010212

Brookhart, S. M. (2008). Feedback that fits *Educational Leadership, 65*(4), 54-59.
https://eric.ed.gov/?id=EJ781221

Cameron, D. (2005). Language, gender, and sexuality: Current issues and new direction. *Applied Linguistics, 26*(4), 482-502.
https://doi.org/10.1093/applin/ami027

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277-299.
https://doi.org/10.1207/s1532799xssr1003_5

Douglas, D. (2010). *Understanding language testing*. Hodder Education.
https://doi.org/10.1177/0265532210373604

Elfi, E. (2019). CALL: The use of Winnerclass Professionals V. 3.0 software in teaching listening. *PROCEEDING IAIN Batusangkar*, *3*(1), 105-110.
https://doi.org/10.4304/jltr.2.5.977-988

Ghahremani, D. (2013). The effects of implementing summative assessment, formative assessment and dynamic assessment on Iranian EFL learners' listening ability and listening strategy use. *Journal of Language and Translation*, *3*(1), 59-68.
http://ttlt.azad.ac.ir/article_514741.html

Godwin-Jones, R. (2011). Emerging technologies autonomous language learning. *Language Learning & Technology, 15*(3), 4-11.

https://scholarspace.manoa.hawaii.edu

Guk, I., & Kellogg, D. (2007). The ZPD and whole class teaching: Teacher-led and student-led interactional mediation of tasks. *Language Teaching Research*, *11*(3), 281-299. https://doi.org/10.1177/1362168807077561

Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (pp. 61-80). Sage.

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge University Press. https://doi.org/10.1177/0734282909347604

Hidri, S. (2014). Developing and evaluating a dynamic assessment of listening comprehension in an EFL context. *Language Testing in Asia*, *4*(1), 1-19. https://doi.org/10.1186/2229-0443-4-4

James, M. (2008). Assessment and learning. In S. Swaffield (Ed.), *Unlocking assessment: Understanding for reflection and application* (pp. 20-35). Routledge.

Jones, M. G., & Brader-Araje, L. (2002). The impact of constructivism on education: Language, discourse, and meaning. *American Communication Journal*, *5*(3), 1-10. https://ac-journal.org/journal/vol5/iss3/special/jones.pdf

Kennedy, K., Chan, J., Fok, P., & Yu, W. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice, 7*(3), 197-207. https://eric.ed.gov/?id=EJ814321

Khoshsima, H., & Mozakka, Z. (2017). The effect of computer-assisted language learning on Iranian upper-intermediate EFL learners' listening skill. *Journal of Applied Linguistics and Language Research*, *4*(2), 81-91. http://www.jallr.com/index.php/JALLR/article/view/520

Lam, R., & Lee, I. (2010). Balancing the dual functions of portfolio assessment. *ELT Journal, 64*(1), 54-64. https://doi.org/10.1093/elt/ccp024

Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, *1*(2), 49-72. https://doi.org/10.1558/japl.v1.i1.49

Lantolf, J. P., & Poehner, M. (2008). Dynamic assessment. In N. Hornberger (Ed.), *The encyclopedia of language and education*: *Language testing and assessment* (pp. 273-285). Cambridge University Press.

Lebedeva, M. Y., Koltakova, E. V., Khaleeva, O. N., & Rusetskaya, M. N. (2016). Computer-assisted language learning for the development of listening skills: A case study of pre-university Russian as a foreign language. *International Journal of Applied Linguistics and English Literature*, *6*(1), 257-265. http://dx.doi.org/10.7575/aiac.ijalel.v.6n.1p.257

Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing, 22*(1), 34-50. http://dx.doi.org/10.1016/j.jslw.2012.11.003

Mashhadi Heidar, D. (2016). ZPD-assisted introduction via web 2.0 and listening comprehension ability. *English for Specific Purposes World, 49*(17) 1-17. http://esp-world.info/Articles_49/Heidar.pdf

Mashhadi Heidar, D., & Afghari, A. (2015). The effect of dynamic assessment in synchronous computer-mediated communication on Iranian EFL learners' listening comprehension ability at upper-intermediate level. *English Language Teaching, 8*(4), 14-23. https://doi.org/10.5539/elt.v8n4p14

McKeough, A., & Lupart, J. L. (2013). *Toward the practice of theory-based instruction: Current cognitive theories and their educational promise*. Lawrence Erlbaum Associates.

Nachoua, H. 2012. Computer-assisted language learning for improving students' listening skill. *Procedia-Social and Behavioral Sciences, 69,* 1150-1159. https://doi.org/10.1016/j.sbspro.2012.12.045

Poehner, M. E. (2005). *Dynamic assessment of oral proficiency among advanced L2 learners of French* [Unpublished doctoral dissertation]. Pennsylvania State University, University Park.

Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, *43*(3), 471-491. https://doi.org/10.1002/j.1545-7249.2009.tb00245.x

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity*, *17*(4), 312-330. https://doi.org/10.1080/10749030903338509

Park, K. (2014). Corpora and language assessment: The state of the art. *Language Assessment Quarterly, 11*(1), 27-44. https://doi.org/10.1080/15434303.2013.872647

Rahimi, M., Kushki, A., & Nassaji, H. (2015). Diagnostic and developmental potentials of dynamic assessment for L2 writing. *Language and Sociocultural Theory*, *2*(2), 185-208. https://doi.org/10.1558/lst.v2i2.25956

Roohani, A., Jam, B., Yeganeh, S., & Domakani, M. R. (2018). The effect of dynamic assessment on L2 learners' listening comprehension. *Bellaterra Journal of Teaching & Learning Language & Literature*, *11*(4), 59-70. https://doi.org/10.5565/rev/jtl3.751

Sehati, S., & Khodabandehlou, M. (2017). Effect of power point enhanced teaching (visual input) on Iranian Intermediate EFL learners' listening comprehension ability. *Journal of Educational Issues*, *3*(2), 29-42. https://doi.org/10.5296/jei.v3i2.12323

Shepard, L. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 1066-1101). American Educational Research Association.

Vahdat, S., & Eidipour, M. (2016). Adopting CALL to improve listening comprehension of Iranian junior high school students. *Theory and Practice in Language Studies*, *6*(8), 1609-1617. http://dx.doi.org/10.17507/tpls.0608.13

Vygotsky, L. S. (1978). Interaction between learning and development. *Readings on the Development of Children*, *23*(3), 29-39. https://oerafrica.org/sites/default/files/L%20&%20L%20reader_section%20one-reading_4.pdf

Vygotsky, L. S. (1998). The problem of age. In R. W. Rieber (Ed.), *The collected works of L. S. Vygotsky* (pp. 187-205). Plenum. http://dx.doi.org/10.1007/978-1-4613-1655-8

Walsh, C. (2001). *Gender and discourse: Language and power in politics, the church and organizations.* Longman.

Zeng, W., Huang, F., Yu, L., & Chen, S. (2018). Towards a learning-oriented assessment to improve students' learning: A critical review of literature. *Educational Assessment, Evaluation and Accountability, 30*(3), 211-250. https://doi.org/10.1007/s11092-018-9281-9

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

The Impact of Teacher-Student Interaction and Academic Self-Concept on EFL Learners' Academic Achievement

Research Article
pp. 225-245

Nafiseh Asadzadeh Maleki[1]
Masoud Zoghi*[2]
Nader Asadi Aidinloo[3]

Received: 2020/12/27 Accepted: 2021/05/12

Abstract
While previous academic achievement studies have investigated a wide range of cognitive variables, certain non-cognitive variables such as teacher-student interaction and academic self-concept in relation to academic success have remained under-researched. To this end, a coefficient correlation research method was used to observe the academic self-concept, and the interaction between teacher and student features which had been identified as having significant roles in English students' academic achievement. The participants of the study consisted of 218 Iranian EFL university learners, with an age range of 18 to 45. Structural equation modeling hypothesis-testing procedures demonstrated that there exists a reasonable fitness between the data and the model ($\chi^2 = 0.37$, GFI = 0.99, CFI = 0.99, RMSEA = 0.01). It was confirmed that the amount of student and teacher relationships meaningfully and directly affected EFL students' academic success ($r = 0.35$, $p = 0.45$). Moreover, academic self-concept was known to be a significant factor in improving students' achievement ($r = 0.24$, $p = 0.32$). The obtained results also indicated the mediating role of learners' academic self-concept in enhancing learners' academic achievement via increasing the interaction between students and their teacher. The fundamental role of psychosomatic features, comprising of students and teacher relations and academic self-concept to foreign language students' educational success was authenticated. The outcomes of this empirical study have some implications for educational settings.

keywords: academic achievement, academic self-concept, correlational method, structural equation modeling, teacher-student interaction

* Corresponding Author
1 Ph.D. Candidate, English Department, Faculty of Humanities, Ahar Branch, Islamic Azad University, Ahar, Iran.nafiseh.asadzadeh@malekaniau.ac.ir
2 Assistant Professor, English Department, Faculty of Humanities, Ahar Branch, Islamic Azad University, Ahar, Iran. m-zoghi@iau-ahar.ac.ir, dr.m.zoghi@gmail.com
3 Assistant Professor, English Department, Faculty of Humanities, Ahar Branch, Islamic Azad University, Ahar, Iran. naderasadi@yahoo.com

DOI: 10.22051/LGHOR.2021.34581.1426

Introduction

Student accomplishment or growth is considered as being of essential concerns for educational institutions that value economic as well as universal progress (Brown & Ryan, 2015). Learners experience different kinds of emotions that are related to educational activities (Raccanello & De Bernardi, 2013). According to (Pekrun, 2006), learners' academic achievement consists of characteristics of students, for instance the notions of self-concept and societal and educational issues. In preceding experiential study (Pekrun & Perry, 2014), the influential role of learners' self-concept in improving their success has been confirmed, while the starring role of the superiority of students and teacher interactions is not fully deliberated. As well, it has been proposed that students' relationship with their teacher significantly improves their success in academic context (Belsky & Pluess, 2009). However, fewer researches have investigated the mediated contribution of learners' academic self-concept in the topic of the association amongst the teacher–student interaction and academic achievement. The present study has two goals: First, to explore the roles of the student and teacher communication and academic self-concept in learners' accomplishment progress. Second: to ascertain the direct and mediated roles of these two variables on learners' academic achievement.

Literature Review

Academic achievement is regarded as a significant principle to evaluate a person's total potentialities and capabilities (Berg & Coetzee, 2014). The concept of academic achievement is well-established as the students' understanding of numerous abilities and educational materials through applying consistent and validated analyses (Joe et al., 2014). Scholars agree that cognitive and non-cognitive skills (Pitt et al., 2014; Veas et al., 2019), such as motivational (Wolff et al., 2018), and situational factors (LoCasale-Crouch et al., 2018) contribute to learners' academic achievement and growth in academic setting. The literature demonstrate a great deal of consideration to the understanding of the ways in which non-cognitive variables such as motivational and situational factors have an effect on students' ability to achieve higher grades (Berg & Coetzee, 2014; Pitt et al., 2012).

The optimistic self-beliefs that concentrates on how healthy, normal, and exceptional people may obtain the thoroughgoing from a lifespan are centered at the core of the positive psychological revolution (Grygiel et al., 2017; Simonsmeier et al., 2020). This positive psychology initiative that concentrates on academic self-concept, draws a vast amount of attention in educational and psychological fields (Marsh et al., 2017). The term self-concept is explained as "a construct of pervasive significance" (Craven, 2008, p. 1) that supports the comprehension of required psychological, behavioral, and educational consequences that reinforce human potential.

The notion of self-concept was primary familiarized by Carl Rogers and Abraham Maslow, addressing how people perceive themselves as pertaining to their capabilities, attitude, values, and uniqueness (Garn & Shen, 2015). In educational settings, self concept is considered as a significant consequence of learning and assists as an indicator of scholastic outcome. It is demonstrated that the notion of self concept is correlated with accomplishment, and academic self□concept is a momentous predictor of professional preference in and out of school (Parker et al., 2015). Khalaila (2015) pointed out that academic self-concept is one of vital key

elements to learners' academic success. He distinguished that the foundation of scholars' academic superiority is the academic achievement in which understanding the various features responsible in predicting, identifying, intervening, or having an effect in academic achievement is critical (Khalaila, 2015).

A number of empirical findings provide numerous illustrations for the constructive weights of a learners' high academic self-concept. According to Harter (2012), academic self-concept (ASC) is defined as an individual estimation of personal cognitive capabilities in academic achievement contexts. The researchers illustrated that high amount of achievement, persistence at the time of failure, interest, and academic choice are all associated with learners' academic self-concept (Gogol et al., 2017; Köller et al., 2006; Marsh et al., 2017). Students who have positive academic self-concept attain advanced educational accomplishment levels and are further successful in professional domains (Marsh et al., 2017; Schneider & Preckel, 2017). In addition to the relationship with performance consequences, constructive academic self-concept is related to other areas of academic success, for instance not as much of university failure (Schiefele & Schaffner, 2015). Accordingly, it is increasingly recognized that a student's positive academic self-concept can simplify a vast amount of educational consequences (Trautwein & Möller, 2016). A plethora of literature have clarified the substantial effect of academic self-concept on achievement (Chen et al., 2013; Marsh et al., 2015). For instance, whenever students are self-confident in their capability to be academically effective, their self-concept will affect their interest, causing them to be more satisfied about learning. At the point when learners reveal passion or enthusiasm for learning they will undoubtedly turn out to be motivated to learn new materials and notions, which will help them to obtain the intentions of pedagogic excellence (Guo et al.,, 2018; Niepel et al., 2014; Pinxten et al., 2014). Teachers as well as parents should establish a strong academic self-concept which would be considered as the first priority over the enhancement of long-term academic achievement (Marsh et al., 2015). Knowledge of students' perceptions of their own academic competencies, feelings, and experiences is noteworthy, since these perceptions will impact how they in turn value academics. Moreover, students' perceptions of the importance of academics will ultimately influence how well or ineffectively they perform academically (Trautwein & Möller, 2016).

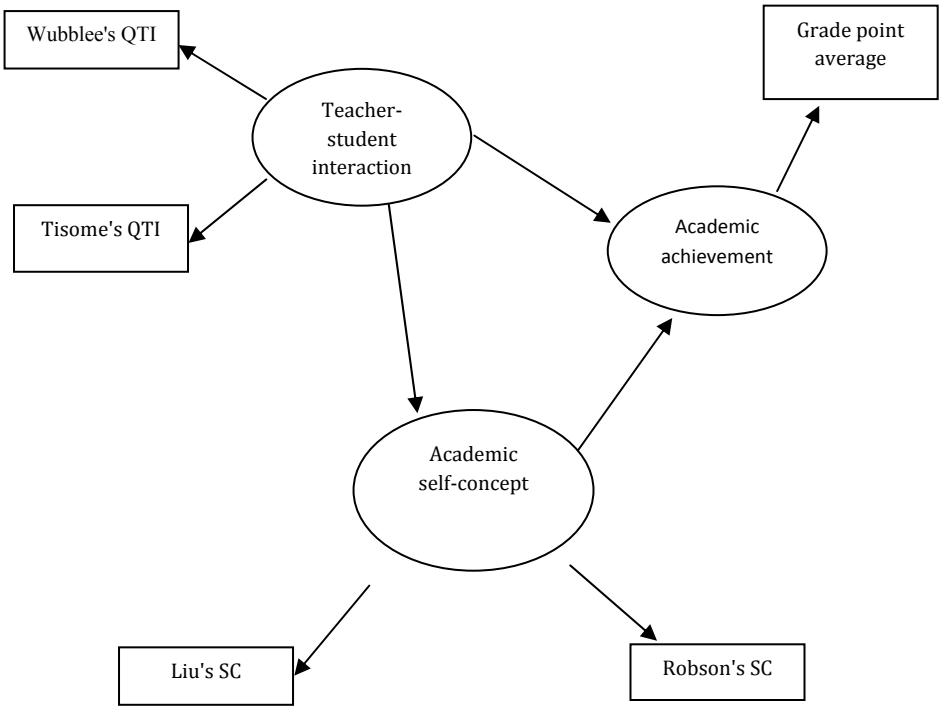
Conversely, in the face of the strong practical provision for the relationships of educational achievement and self-concept, rare studies have considered the mediated role of this concept concerning the relationship of academic achievement and the amount of teacher-student contact. As an alternative, most of the existing achievement motivation studies have relied mainly on the direct consequence of academic self-concept on achievement, even though the ASC's mediated role might be the most important variable in the establishment of learners' strong achievement (Guo et al., 2018) and fast progress in cognitive and academic skills (Harter, 2012). Some researchers (e.g. Cadima et al., 2010) elucidated that educators' interaction quality with students in the class is increasingly acknowledged as of having dominant significance for learner achievement in classroom (Cadima et al., 2010). The relationship between educator and students is observed as the utmost influential feature in educational situation. This sort of relation that is considered as academic motivation and school engagement, expresses the basis of the social context in which learning takes place (Clem et al., 2020; Pekrun & Perry, 2014). A number of variables such as gender have an influential effect on the interactions

between teacher and students which in turn influence a learner's academic outcomes and performance. Hughes & Chen (2011) consider that supportive and positive relations that occur between students and teachers ultimately stimulate a "sense of school belonging" and inspire learners to "participate cooperatively in classroom activities" (p. 278).

Generally speaking, teacher–student interactions are considered at two levels, the classroom level and the dyadic level. An attachment-based perspective is frequently utilized in educational researches (Pekrun & Perry, 2014; Verschueren & Koomen, 2012). This perspective concentrates on the emotional constituents of the dyadic relations that exist among the students and teachers. More explicitly, it focuses on three major components, namely dependency, closeness, and conflict, in the relationship between the educator and a specific learner (Verschueren, 2015). Closeness discusses the amount of warmth, open communication between the student and instructor, and affection, while conflict encompasses the amount of undesirable affect communicated and experienced in the interaction (Ahnert et al., 2012). Dependency speaks about the student's insecure behavior toward the instructor and the high degree of reliance on the teacher (Ahnert et al., 2012; Vervoort et al., 2015). Nevertheless, these dyadic relationships are surrounded by the wide-ranging classroom environment, which have an impact on learner's progress in its own manner (Howes et al., 2011). Teacher–student interactions are increasingly noteworthy for the adjustments of students' socioemotional features and their success in educational settings (Lei et al., 2018; Pakarinen et al., 2018). Numerous theoretical explanations have been proposed as to why the quality of teacher–student interaction is expected to similarly be associated with learners' achievement. For instance, supportive and close teacher–student interactions can improve learners' common sense of belonging and social relatedness within the class, which, consecutively, may possibly stimulate positive and impede negative feelings towards learning (Clem et al., 2020; Lei et al., 2018; Skinner et al., 2014).

Furthermore, whereas empirical studies have noticeably determined the significance of societal, temporal and dimensional contrasts for academic self-concept construction, rare researches have experimentally examined their joint influence as mediator in learners' achievement and interaction relationship (Wolff et al., 2018). As such, in spite of an overabundance of self-concept researches, most of the individual studies focused only on the reciprocal effects of ASC and achievement, in actual fact none have exclusively juxtaposed the contributive role of this notion in relation to the developmental perspective of achievement. In this context, the present paper designed to rigorously test: a) the developmental interplay between teacher-student interaction and a standardized academic achievement measure at university level; b) the actual magnitude of the special impressions of academic self-concept on EFL university students' success; c) the magnitude to which academic self-concept play a mediator role in learners' academic achievement, respectively. Data was collected from among university students. Following Marsh and Martin (2011), structural equation modeling (SEM) was utilized in this study to account for measurement error and so was anticipated to produce supplementary precise parameter assessments (Marsh et al., 1998) (Figure 1).

Figure 1
The Hypothesized SEM Model



The Current Study

In the mainstream of researches discovered heretofore, it was demonstrated that academic achievement is a major determinant in the increase concept model. According to this model, self-concept tends to enhance on the involvement in explicit and implicit trainings of the scholars (Ahmed & Bruinsma, 2006). As an indicator of accomplishment in any systematic attempt, student's success and other variables have an effect in any educational system in which this concerns persuades and attract numerous researchers. Generally, it is still open whether the different models for considering the association between self-concept and achievement obtain comparable consequences. As well, the teacher-student relationship was presumed to be predicator and academic achievements were considered as an outcome, thus it is commonsense to hypothesize that academic outcomes are influenced by interactions with teachers (Goetz et al., 2013; Westphal et al., 2018). By contrast, almost nothing is known about whether learners' academic self-concept have a mediator role in achievement-related outcomes and teacher-student interaction relationship, even though there exist good explanations to express that this might be the case (Nelson & Schunn, 2009). The current study aimed to fill the aforementioned gap by scrutinizing whether academic self-concept mediate the association between learners' academic success and the extent of their interaction with teacher. In this research paper facets of expectancy value theory (Eccles & Wigfield, 1995) and self-determination theory (Deci & Ryan, 1985) were combined, to expand our knowledge of the notions of academic achievement, academic self-

concept, and teacher-student interaction along with the relationships that exist among them. More specifically, this paper attempted to scrutinize effects of the research variables on improving EFL learners' academic achievement. The study was conducted to investigate a series of research questions, as follows:

- *Is there a direct relationship between teacher-student interaction and EFL students' success?*
- *Does EFL students' self-concept have a direct effect on EFL students' success?*
- *Is the effect of teacher-student interaction on EFL learners' success mediated by self-concept?*

Materials and Methods

Design of the Study

This investigation was an attempt to survey the relative contribution of the amount of teacher-student relations and academic self-concept to EFL learners' success. The study utilized the descriptive correlational survey research design. Parametric statistical tools were employed to determine students' level of self–concept, their achievement in academics, the amount of teacher- student interaction, and the direct and indirect relationships that exist among these variables (Creswell, 2008). Likewise, the present research was a cross sectional design that involves collecting of the required information from the sample group on one occasion. In this research method, the investigator is able to infer and generalize required data in relation to a larger population according to the obtained data of smaller sample group (Creswell, 2008).

Participants

For the purpose of inspecting the contributive role of teacher-student interaction, self-concept to learners' academic success, a-priori Sample Size Calculator for SEM was utilized. The sample size required for this study that drew on the SEM was calculated to be 218 EFL learners. Accordingly, a total number of 218 Intermediate EFL learners varying in age from 18 to 45 were participated in this study. The participants were EFL learners in Islamic Azad University, Tabriz Branch, Iran. Of the entire number of these 218 students, 46.8 % were male, and 53.2 % were female, with Turkish and Farsi as their native language.

Procedure

Once the research framework was established, the required data to conduct this research were gathered throughout the educational year 2017. The researchers secured the appropriate permissions from the aforementioned university's administrative offices. At the start of the study, the researchers personally visited the learners at the university and briefed the participants about the resolution and the prominence of the study. As such, we explained the research procedure and the methods in which they were supposed to respond the questions of each questionnaire. The students were required to provide their own opinion for each question as stated by the determined instruction of each questionnaire. With the aim of getting high response and authentic rate, the researchers read each question out loud and the students were asked to choose one answer of their own desired reply in the questionnaires. During the procedure, the students were assured that their opinions would not be shared with anyone. They were just demanded to exactly

provide their biographical data containing their age, sexual category, score, and their grade point average (GPA). Since, the learners' success was assessed in line with the students' grade point average score, we emphasized precise self-reporting of their GPA. Every questionnaire was coded by the computer, and then two statistical software, named SPSS 22.0 and LISREL 8.80, were utilized to examine the proposed model.

Measures

To answer research questions, the following instruments were utilized:

Self-Concept Ability: learners' academic self-concept ability was assessed via utilizing two questionnaires: a) Robson's (1986) self-concept questionnaire – this 7-item questionnaire measures five sub-categories of self-concept, i.e., attractiveness and approval by others, contentment and worthiness, confidence value of existence, determinism and significance, and resilience. The Cronbach's alpha reliability measure for the self-concept questionnaire was calculated to be 0.80; and b) Liu and Wang's (2008) Academic Self-concept Questionnaire (ASCQ) which aims to assess the emotional states and insights of the students' academic competence. Twenty items are included in the questionnaire with 6-Likert scale (yes, no, no always, no sometimes, yes sometimes, yes always). Two major subscales measured in this questionnaire are confidence and effort. This questionnaire was, also, evaluated in terms of its reliability and the results indicated the reliability measure of .92. An example of a sample item of this questionnaire is as follows:

Question	*Yes*	*No*	*No always*	*No sometimes*	*Yes sometimes*	*Yes always*
I can follow the lesson easily	.....	....	……	……	.....	.....

Teacher-Student Interaction: The quality and the features of the learners' relations with their instructors were measured through the use of two different questionnaires, too. a) Questionnaire on Teacher Interaction (QTI): The scale was designed to measure learners' experience of leadership, strict, uncertain, responsibility/freedom, helping/friendly, understanding, dissatisfied, and admonishing on a Likert scale, ranges from letter A (never, A stands for 1) to letter E (always, E stands for 5). The amount of Cronbach's alpha for the feature of closeness was demonstrated to be 0.82, and for conflict, it was 0.76. The Cronbach's alpha reliability of this questionnaire was indicated to be agreeable (r = 0.81). b) a self-reporting Questionnaire on teacher-student interaction (QTI) modified by Nugent (2008). The questionnaire includes 48 items about the students' performance in the class and the numbers range from 1 to 5 (1 = never, and 5 = always). The reliability of the instrument was calculated as 0.80. The following sample item is an example of this questionnaire:

1. This instructor talks enthusiastically about her/his subject. 1 2 3 4 5

Academic Achievement: Students' grade point averages (GPA) were utilized to exhibit the learners' academic success, and were observed to serve as measurements of the English achievement achieved at the university level. EFL learners were required to self-report their grade point average score which ranges from 0 to 20.

Data Analysis

To investigate the research questions, the various questionnaires were submitted to the students and required data were collected. The collected data were coded and scrutinized via the use of the program Statistical Package for the Social Sciences (SPSS) version 22.0. At the outset, the demographic data of the students were analyzed by the use of statistical analysis to scrutinize the descriptive data involving of the standard deviations, the mean, and the standard error of means. Moreover, the coded data in SPSS were imported into LISREL Software version 8.80. Then, the researchers applied Confirmatory Factor Analysis (CFA) to determine the relationship among latent and observable variables. The foremost intention of confirmatory factor analysis is to check the fitness of the data to a theorized measurement model through the use of different indices of model fit index. In Model fit the researchers aims to explore the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the Tucker–Lewis index (TLI). A model characteristically fits the data while the value of p is not significant. Whenever the value of RMSEA is lower than 0.06 and SRMR values is under 0.08, it is demonstrated that there exists a comparatively good fitness between the theorized model and the obtained data (Hu & Bentler, 1999). Because of the sensitivity of the $\chi 2$ test to sample size, applying relative goodness-of-fit indices is correspondingly suggested in large sample size cases (Bentler & Bonett, 1980). At last, Structural equation modeling (SEM) of the research variables was provided using LISREL software. Structural equation modeling is defined as a statistical analysis procedure that is utilized to evaluate the structural associations.

Results

To begin with, the researchers calculated Skewness and Kurtosis of the research variables along with the correlational association of these variables. In order to indicating the data normal distribution, kurtosis and skewness of the scores are revealed in Table 1.

Table 1
Skew and Kurtosis of Research Variables

Variable	Mean	Standard Deviation	Skewness	Kurtosis
GPA	17.34	1.64	-0.59	-0.40
Robson's SC	53.20	12.29	0.19	-0.87
Liu's SC	51.02	11.62	-0.42	0.30
Wubble's QTI	212.97	40.09	-0.59	0.76
Tisome Nugent's QTI	162.32	31.82	-0.21	0.68

As it is demonstrated in Table 1, the distributions of the academic success showed negative skewness (-0.59) and negative Kurtosis (-0.40). The skewness of academic self-concept variables ranged from 0.19 to -0.42 and the kurtosis of these variables was between -0.87 and 0.30. Correspondingly, the distributions of teacher-student interaction variables authorized positive kurtosis between 0.76 and 0.68 as well as negative skew between -0.59 and -0.21. Generally speaking, the absolute skewness value of more than 2 and the absolute kurtosis value of over 7 are demonstrated to be unacceptable (Finney & DiStefano, 2006). In the present paper, the kurtosis and skewness of the distributions of research variables were surrounded by the range of acceptability and normality. Table 2 below provides information on

the correlational statistics that exist among the study variables.

Table 2
Correlational Statistics among Variables

No.	Variable	1	4	5	7	8
1	GPA	1				
4	Robson's SC	0.31**	1			
5	Liu's SC	0.32**	0.47**	1		
7	Wubble's QTI	0.45**	0.40**	0.39**	1	
8	Tisome Nugent's QTI	0.41**	0.35**	0.31**	0.71**	1

** $p< 0.01$

As demonstrated in Table 2, there were significant average correlations among the research variables ($p < 0.01$).

In the current study, the researchers applied Confirmatory factor analysis to evaluate the rationality of measures. This research paper used indicies from different classifications to evaluate goodness of data-model fit, to be precise, camparative fit indices, absolute fit indices, and parsimonious fit indices (Hu & Bentler, 1999). The data-model fit indices are: a) Comparitive Fit Index (CFI), Normed Fit Index (NFI), and Non-Normed Fit Index (NNFI) as comparitive fit indices, b) Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), and Standardized Root Mean Square Residual (SRMR) as absolute fit indices, and c) Chi-Square/Degree of Freedom ($\chi 2$/df), Parsimonu Normed Fit Index (PNFI), and Root Mean Square Error of Approximation (RMSEA) as parsimonious fit indices. Table 3 illustates the fitness of the proposed structural equation modeling.

Table 3
Data-Model Fit Indices

	Absolute fit indices		
SRMR	AGFI	GFI	Index
0.01	0.99	0.99	Value
<0.05	>0.80	>0.90	Acceptable range
	Comparative fit indices		
NNFI	NFI	CFI	Index
0.99	0.99	0.99	Value
>0.90	>0.90	>0.90	Acceptable range
	Parsimonious fit indices		
RMSEA	PNFI	X2/df	Index
0.01	0.61	0.37	Value
<0.08	>0.60	<3	Acceptable range

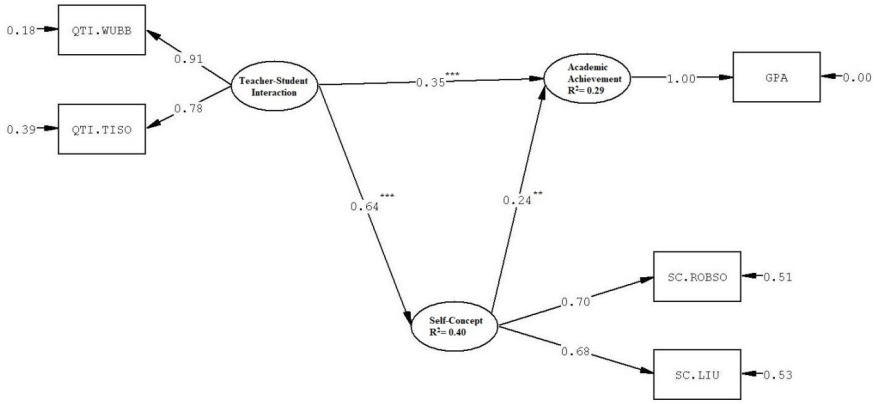
The goodness of fit index (GFI) of the proposed model was 0.99 and the adjusted goodness of fit index (AGFI) was 0.99. Based on the obtained outcomes, the Standardized Root Mean Square Residual (SRMR) equaled 0.01. The entire measures of absolute fit indices are within the accepted ranges (GFI = 0.90, AGFI = 0.08, and SRMR = 0.05), representing thet the hypothetical model was a good fit of the data (Schumacker & Lomax, 2010). The normed fit index (NFI) of the model calculated to be 0.99, the comparitive fit index (CFI) of the operationalized model

was 0.99 and the root mean square error of approximation (RMSEA) equaled 0.01, and non-normed fit index (NNFI) signifying the value of 0.99 is within the significant value of 0.90. The value of CFI (0.99) falls within the meaningful range of 0.90 or higher, designating this model was a good fit of the data (Schumacker & Lomax, 2010). The $\chi 2$ ratio that is divided by degree of freedom was equal to 0.37. Consistent with Carmines and McIver (1981), a $\chi 2$ ratio of less than 2.00 advocates a good fit, consequently approving that this model fits the data.

As aforementioned, the researchers employed LISREL software (version 8.80) to survey factor loadings of the research variables on the latent construct and coefficients for the constructs' directional paths. In path analysis, the pathways are characterized as equations and exemplified as path diagrams. Figure 2 exhibits the structural equation model of the study.

Figure 2

Structural Equation Modeling of the Study



Correspondingly, the direct and indirect effects of variables are revealed in Table 4.

Table 4

The Direct and Indirect Effects of Each Variable on Academic Achievement

Path	Direct effect	Indirect effect	Total effect	R^2
To Academic achievement from				0.29
Academic Self-concept	0.24^{**}	-	0.24^{**}	
Teacher-student interaction	0.35^{***}	0.15^{**}	0.50^{***}	
To Academic Self-concept from				0.40
Teacher-student interaction	0.64^{***}	-	0.64^{***}	

**p < 0.01, ***p < 0.001

As established in Table 4, learners' academic self-concept had a significant prognostic influence on their educational achievement ($\beta = 0.24$, $p < 0.01$). The amonunt of the relatioship between educator and learner, harmoniously, had a momentous and significant prognostic role in signifying academic success ($\beta = 0.35$, $p < 0.001$). The findings provided the required data for confirming the first two research questions, appealing that both teacher-student relations and academic self-

concept definitely influenced students' academic succes. Additionally, the indirect effect of teacher-student relations via academic self-concept on academic achievement was designated to be (0.15), which confirmed to be noticeably significant in line with standardized significant level ($p < 0.01$). Accordingly, it is rational to claim that students' academic self-concept had a mediating role regarding teacher-student relations and academic success relationship. That is to say that improving the quantity of the relationship between teacher and students caused the enhancement in students' academic self-concept which resulted in the expansion of academic success.

Discussion and Conclusion

The research variables multifaceted nature that were influential in the learners' academic procedure entails the prerequisite to consider their diverse levels of relations, since this is an imperative phase in which motivational, contextual, and cognitive mechanisims can be incorporated as noticeable features in improvement. The current research extended upon preceding studies on teacher-student interactioin, academic self-concept, and academic success (Goetz et al., 2013; Goetz et al., 2012; Westphal et al., 2018) by displaying that learners' academic self-concept moderated the association between learners' academic success and the amount of the interaction with their teacher. The primary resolution of this paper was to explore the direct effect of learners'acadademic self-concept on their academic growth. Correspondingly, the hypothesized relatioship between teacher-student interaction and students' academic achievement was examined. As a final point, the study attempted to scrutinize the contributive role of academic self-concept and teacher-student relations in EFL students' academic success. Consistent with the previous studies and expectations on direct effect of teacher-student interaction on achivement (hypothesis 1), learners' academic success was established to be directly influenced by the quality of teacher-student communication. This authorized the prospect that learners learn more when they positively communicate with their teacher and their classmates (Abrantes et al., 2007; Kiuru et al., 2020). Learners better understand the knowledge and become more committed to learn more and more through interacting with others (Hay et al., 2004). The results for estemating academic achievement according to the amount and the quality of teacher-student interaction demonstrated that a students reached advanced levels of achievement and lower levels of tediousness and anixty when their interaction with the teacher was closer and less conflictual, whereas conflictual teacher–student association was correlated with higher anxiety levels. As theorized, the obtained results of the study are in acordance with preceding researches indicating that the teacher and students relationships' quality has a noteworthy effect on student academic performance (Frenzel et al., 2007; Goetz et al., 2013; Lei et al., 2018; Skinner et al., 2014). Likewise, the current research results are consistent with the outcomes of many preceding investigations which made significant contribution to the well-established body of investigation on the importance of teacher interaction in the class (Arbeau et al., 2010; Elmelid et al., 2015; Goetz et al., 2013; Jellesma et al., 2015). Findings support the view of variouse researchers (Kenny & Adriana, 2009; Košir & Tement, 2014; Kyriakides, 2005; Wubbels & Brekelmans, 2005) that, in fact, learner perceptions of teacher interaction have a remarkable role in student success and can be utilized as a predicator for student achievement. The consequences are as well in accordanace with previouse studies indicating that caring and supportive

interpersonal relationships among teachers and students in school reveal supplementary positive attitudes and values and more satisfaction with education. This authorizes the belief that students learn more when they positively communicate with their teacher (Bakadorova & Raufelder, 2016; Clem et al., 2020; Sakiz, 2012). It is conceivable that in a supportive teacher–student relationship, learner' rudimentary psychosomatic requirements for relatedness are satisfied and teachers are more sensitive and responsive to students' individual learning-related desires. This may allow learners to involve in educational tasks with satisfaction and interest (Skinner et al., 2014).

In the previouse studies on self-concept impacts on achievement, varoiuse consequences have been stated with reference to the direction and the extent of effects between learners' academic self-concept and their academic success. Nevertheless, as expected in the second hypothesis, the results of the current study evidenced that academic self-concept has a predicted role in learner academic achivement. This results corraborates other experiential conclusions in this domain (Ganley & Lubienski, 2016; Gujare & Tiwari, 2016; Marsh et al., 2015; Marsh & Martin, 2011; Niepel et al., 2014; Pinxten et al., 2014). The ontained results agrees to the notion that learners' academic self- concept is a more concrete and noticeable variable for studenets depelopment and, consequently, impact their achievement during scholastic tasks. Considering the prominant influence of academic self-concept in enhancing academic progress, Grygiel et al. (2017) propsed that self concept is a contributing factor of academic success (Grygiel et al., 2017). Many researchers have supported the contention that there exist a constructive association between students academic self-concept and performane in aquiring a language (Hoferichter et al., 2018; Kulakow, 2020; Rady et al.,) that investigated the associations between learners' performance and their academic self-concept. The consequences authorized the existence of a momentous statistical correlation between students' educational performance and their academic self-concept. The learners who have high level of academic accomplishment receive reward and have more opportunities in future.

The consequences of the existing study similarly advocate the third research hypothesis which claims that academic self-concept mediates the association between the students' academic performance and the amount of teacher-student contact. To our knowledge, fewer studies are known to explore the associations between the achievement and interaction that is mediated by the role of academic self-concept. The obtained consequences of the study indicated that academic self-concept had a noteworthy relationship with academic success, which is a confirmed to be a fundamental mediator in the interply between teacher-student interaction and academic success in previouse socio-cognitive principles. These consequences strengthen the reliable sound effects attributable to the interaction among students and teacher and the self-concept that is essential to be regarded in existing achievement models, according to a interactional framework (Greenwald et al., 2002). Likewise, the prominence of self-concept as a mediator is in accordance with earlier researches (Díaz et al., 2016; Green et al., 2012; Pinxten et al., 2013; Rady et al., 2016; Sabol & Pianta, 2012) which established it as a fundemental factor to academic achievement. According to the consequences of the current study and abovementioned data, the mediating consequence of learners' academic self-concept in the relationship between teacher-student interaction and academic achievement allows us to confirm significant implications. Societal communications among

teachers and thier students are considered as essential variable in academic success, since they can result in enhanced emotional support and prosocial behaviors (Kraft & Rogers, 2015; Pino-Pasternak et al., 2010). The current outcomes provided supplementary confirmation for the mediated role of academic self-concept regarding teacher-student interaction and academic achievement which are in accordance with previouse studies (Gutiérrez et al., 2017; McFarland et al., 2016; Veas et al., 2019; Verschueren et al., 2012). Generally, the study confirmed the significant relationship between the amount and the quality of teacher-student interaction and learners' academic development, indicating that as the students privided with the opportuinity to interact sincerely and warmly with their teachers, they'll be more successful in educational domains. Furthermore, the research outcomes revealed a closer association between academic self-concept and academic performance measures. In the mainstream of research findings reveled heretofore, it would be sensible to declare that academic self-concept strongly and definitely foretells learners' academic success and mediates the relationsip between teacher-student interaction and academic success. In this sense, teachers and students formal and informal interactions entail worthy levels of societal self-support, given constant feedback, which can strengthen learners' self-concept to accomplish achievement (Díaz et al., 2016; Gutiérrez et al., 2017).

One limitation of the present study is that, rather than longitudinal studies, this study is grounded on a cross-sectional research design. Longitudinal designs permit the investigator to highlight the probable causativeness, along with trends and designs respecting to alterations in learners' beliefs and attitudes during education and may provide responses several questions with reference to behaviors related to academic achievement. One more limitation to this research is that it makes use of only self-reported data. Several detriments are accompanied with the use of self-completion questionnaires. Finally, the survey instruments, although developed from previously-tested and reliable resources and research, has never been tested in a study.

The exceptional theoretical and methodological attitudes to this study created numerous major theoretical implications on the subject of above-mentioned variables. Considering various disciplinary viewpoints and paradigms, the theoretical framework administered in this study characterizes the primary endeavor to incorporate the entire range of academic achievement and teacher-student interaction. Numerous prevoiuse studies have reliably replicated the benefits of constructive teacher-student relations in enhancing amended student consequences. Improving a positive teacher-student relationship should be one of numerous objectives to assist develop new teachers and preserve effective tenured teachers. A very significant section of this study is the fact that information is available on what constitutes "quality teaching" (Wubbles & Levy, 1993). The teachers who are best in their profiesion are very strong in controlling their clssroom and have sincer and friendly relations with the students, therefore, they are less dissatisfied, critical, and uncertain. On the whole, best teachers are extremely cooperative, supportive, and dominant (den Brok et al., 2002). Teachers play pivotal roles facilitating classroom instructions. As such they should adopt a variety of effective practices to successfully promote both students' academic achievement and the degree of relationship with learners. The structural equation modeling applied in the current study offers indication of teacher-student interaction influence on student accomplishment. The consequences of the study make available a robust

consideration of the procedures that generate students' educational success by signifying that student achievement necessitates a mishmash of teacher's sincer and support.

On the topic of the limitations of this research, forthcoming surveys should reproduce the outcomes originated in this research to provide investigators and instructors with more assurance in the meaningful sound effects of close teacher-student relations, self-regulation, and interest on student academic performance. In this research paper, the non-cognitive variables were measured using different questionnaires and academic achievement was evaluated by the use of learners' grade point average. Furthermore, although this was a cross-sectional study, the effect of aforementioned variables on student academic achievement also needs to be examined in a longitudinal study. Alternative domain of forthcoming studies could take account of specific sub-groups, for instance, sexual category and social and economic position to attempt to diminish variables' variability. The findings of this study could be beneficial for curriculum development and instructional design. Curricula could be developed that would comprise instructional components in domains which are indicated to considerably impact learners' language performance. The outcomes of the study could also notify language learner centers and serve as an evaluation of their efficiency. There has been an argument about the significant academic performance and sympathetic and mediated contribution of teacher-student interaction and academic self-concept to achieve best student consequences. The conclusions of the current study reinforced the benefits of these variables in augmenting students' achievement. It was correspondingly found that the effects of the relationship between teacher and student on academic progress was mediated by the concept of academic self-concept.

References

Abrantes, J. L., Seabra, C., & Lages, L. F. (2007). Pedagogical affect, student interest, and learning performance. *Journal of Business Research, 60*(9), 960-964. https://doi/org/10.1016/j.jbusres.2006.10.026

Ahmed, W., & Bruinsma, M. (2006). A structural model of self-concept, autonomous motivation and academic performance in cross-cultural perspective. *Electronic Journal of Research in Educational Psychology,4*(3),551-576. http://dx.doi.org/10.25115/ejrep.v4i10.1214

Ahnert, L., Harwardt-Heinecke, E., Kappler, G., Eckstein-Madry, T., & Milatz, A. (2012). Student–teacher relationships and classroom climate in first grade: How do they relate to students' stress regulation? *Attachment & human development, 14*(3), 249-263. https://doi/org/10.1080/14616734.2012.673277

Arbeau, K. A., Coplan, R. J., & Weeks, M. (2010). Shyness, teacher-child relationships, and socio-emotional adjustment in grade 1. *International Journal of Behavioral Development, 34*(3), 259-269. https://doi.org/10.1177/0165025409350959

Bakadorova, O., & Raufelder, D. (2016). Do socio-motivational relationships predict achievement motivation in adolescents with high and low school self-concepts? *The Journal of Educational Research, 109*(3), 219-231. https://doi.org/10.1080/00220671.2014.942031

Belsky, J., & Pluess, M. (2009). The nature (and nurture?) of plasticity in early human development. *Perspectives on Psychological Science, 4*(4), 345-351. https://doi/org/10.1111/j.1745-6924.2009.01136.x

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin, 88*(3), 588-604. https://doi/org/10.1037/0033-2909.88.3.588

Berg, G. V. D, & Coetzee, L. R. (2014). Academic self-concept and motivation as predictors of academic achievement. *International Journal of Educational Sciences, 6*(3), 469-478. https://doi.org/10.1080/09751122.2014.11890158

Brown, K. W. & Ryan, R. M. (2015). A self-determination theory perspective on fostering healthy self-regulation from within and without. In S. Joseph (Ed.), *Positive psychology in practice,* (pp.139-157). https://doi.org/10.1002/9781118996874.ch9

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher-student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of School Psychology, 48*(6), 457-482. https://doi/org/10.1016/j.jsp.2010.09.001

Craven, R. (2008). The centrality of the self-concept construct for psychological wellbeing and unlocking human potential: Implicat. *Educ. Child Psychol*, *25*, 104-118. https://d1wqtxts1xzle7.cloudfront.net/3624037/cravenmarsh-libre.pdf

Chen, S. K., Yeh, Y.C., H, F.-M., & Lin, S. S. (2013). The relationship between academic self-concept and achievement: A multicohort-multioccasion study. *Learning and Individual Differences, 23*, 172-178. https://doi.org/10.1016/j.lindif.2012.07.021

Clem, A. L., Rudasill, K. M., Hirvonen, R., Aunola, K., & Kiuru, N. (2020). The roles of teacher-student relationship quality and self-concept of ability in adolescents' achievement emotions: Temperament as a moderator. *European Journal of Psychology of Education, 36*, 263-286. https://doi.org/10.1007/s10212-020-00473-6

Craven, R. G., & Marsh, H. W. (2008). The centrality of the self-concept construct for psychological wellbeing and unlocking human potential: Implications for child and educational psychologists. *Educational and Child Psychology*, *25*(2), 104-118.

Creswell, J. W. (2008). *The selection of a research design. Research design: Qualitative, quantitative, and mixed method approaches* (4th ed). SAGE Publications.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality, 19*(2), 109-134. https://doi.org/10.1016/0092-6566(85)90023-6

den Brok, P., Brekelmans, M., Levy, J., & Wubbels, T. (2002). Diagnosing and improving the quality of teachers' interpersonal behaviour. *International Journal of Educational Management, 16*(4), 176-184. https://doi/org/10.1108/09513540210432155

Díaz, E. R., Fernández, A. R., Zabala, A. F., Revuelta, L. R., & Rey-Baltar, A. Z. (2016). Apoyo social percibido, autoconcepto e implicación escolar de estudiantes adolescentes. *Revista de Psicodidáctica, 21*(2), 339-356. https://doi/org/10.1387/RevPsicodidact.14848

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and social psychology bulletin*, *21*(3), 215-225. https://doi.org/10.1177/0146167295213003

Elmelid, A., Stickley, A., Lindblad, F., Schwab-Stone, M., Henrich, C. C., & Ruchkin, V. (2015). Depressive symptoms, anxiety and academic motivation in youth: Do schools and families make a difference? *Journal of Adolescence, 45*(2), 174-182. https://doi/org /10.1016/j.adolescence.2015.08.003

Finney, S.J. & DiStefano, C. (2006). Non-normal and Categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*, (pp. 269-314). Information Age Publishing.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*(5), 478-493. https://doi.org/10.1016/j.learninstruc.2007.09.001

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences, 47*(2), 82-193. https://doi.org/10.1016/j.lindif.2016.01.002

Garn, A., & Shen, B. (2015). Physical self-concept and basic psychological needs in exercise: Are there reciprocal effects? *International Journal of Sport and Exercise Psychology, 13*(2), 169-181. https://doi.org/10.1080/1612197X.2014.940994

Goetz, T., Lüdtke, O., Nett, U. E., Keller, M. M., & Lipnevich, A. A. (2013). Characteristics of teaching and students' emotions in the classroom: Investigating differences across domains. *Contemporary Educational Psychology, 38*(4), 383-394. https://doi.org/10.1016/j.cedpsych.2013.08.001

Goetz, T., Nett, U. E., Martiny, S. E., Hall, N. C., Pekrun, R., Dettmers, S., & Trautwein, U. (2012). Students' emotions during homework: Structures, self-concept antecedents, and achievement outcomes. *Learning and Individual Differences, 22*(2), 225-234. https://doi.org/10.1016/j.lindif.2011.04.006

Gogol, K., Brunner, M., Martin, R., Preckel, F., & Goetz, T. (2017). Affect and motivation within and between school subjects: Development and validation of an integrative structural model of academic self-concept, interest, and anxiety. *Contemporary Educational Psychology, 49*(3), 46-65. https://doi.org/10.1016/j.cedpsych.2016.11.003

Green, J., Liem, G. A. D., Martin, A. J., Colmar, S., Marsh, H. W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of Adolescence, 35*(5), 1111-1122. https://doi.org/10.1016/j.adolescence.2012.02.016

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*(1), 3-25. https://doi/org /10.1037/0033-295x.109.1.3

Grygiel, P., Modzelewski, M., & Pisarek, J. (2017). Academic self-concept and achievement in Polish primary schools: Cross-lagged modelling and gender-specific effects. *European Journal of Psychology of Education, 32*(3), 407-429. https://doi.org/10.1007/s10212-016-0300-2

Gujare, S. K., & Tiwari, G. K. (2016). Academic self-concept and academic outcome of the graduate students: The mediating role of socioeconomic status and gender. *International Journal of Education and Psychological Research, 4*(5), 1-7. https://www.researchgate.net/publication/310461860

Guo, J., Marsh, H. W., Parker, P. D., & Dicke, T. (2018). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learning and Instruction, 58*(2), 210-219. https://doi.org/10.1016/j.learninstruc.2018.07.007

Gutiérrez, M., Tomás, J.-M., Romero, I., & Barrica, J.-M. (2017). Perceived social support, school engagement and satisfaction with school. *Revista de Psicodidáctica (English ed.), 22*(2), 111-117. https://doi.org/10.1016/j.psicoe.2017.05.001

Harter, S. (2012). *The construction of the self: Developmental and socio-cultural foundations.* Guilford Press.

Hay, A., Hodgkinson, M., Peltier, J. W., & Drago, W. A. (2004). Interaction and virtual learning. *Strategic change, 13*(4), 193-204. https://doi.org/10.1002/jsc.679

Hoferichter, F., Lätsch, A., Lazarides, R., & Raufelder, D. (2018). The big-fish-little-pond effect on the four facets of academic self-concept. *Frontiers in Psychology, 9*(2), 1247-1256. https://doi/org /10.3389/fpsyg.2018.01247

Howes, C., Guerra, A. W., Fuligni, A., Zucker, E., Lee, L., Obregon, N. B., & Spivak, A. (2011). Classroom dimensions predict early peer interaction when children are diverse in ethnicity, race, and home language. *Early Childhood Research Quarterly, 26*(4), 399-408. https://doi.org/10.1016/j.ecresq.2011.02.004

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Jellesma, F. C., Zee, M., & Koomen, H. M. (2015). Children's perceptions of the relationship with the teacher: Associations with appraisals and internalizing problems in middle childhood. *Journal of Applied Developmental Psychology, 36*(3), 30-38. https://doi.org/10.1016/j.appdev.2014.09.002

Joe, A. I., Kpolovie, P. J., Osonwa, K. E., & Iderima, C. E. (2014). Modes of admission and academic performance in Nigerian Universities. *Merit Research Journal of Education and Review, 2*(9), 203-230. http://www.meritresearchjournals.org/er/index.htm

Kenny, M. C., & Adriana, M. (2009). Children's self-concept: A multicultural comparison. *Professional School Counseling, 12*(3). https://doi.org/10.1177/2156759X0901200306

Khalaila, R. (2015). The relationship between academic self-concept, intrinsic motivation, test anxiety, and academic achievement among nursing students: Mediating and moderating effects. *Nurse Education Today, 35*(3), 432-438. https://doi/org /10.1016/j.nedt.2014.11.001

Kiuru, N., Spinath, B., Clem, A. L., Eklund, K., Ahonen, T., & Hirvonen, R. (2020). The dynamics of motivation, emotion, and task performance in simulated achievement situations. *Learning and Individual Differences, 80*(2),73-86. https://doi.org/10.1016/j.lindif.2020.101873

Köller, O., Trautwein, U., Lüdtke, O., & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe [On the interplay of academic achievement, self-concept, and interest in upper secondary schools]. *Eitschrift für Pädagogische Psychologie, 20*(1/2), 27-39. https://doi.org/10.1024/1010-0652.20.12.27.

Košir, K., & Tement, S. (2014). Teacher-student relationship and academic achievement: A cross-lagged longitudinal study on three different age groups. *European Journal of Psychology of Education, 29*(3), 409-428. https://doi.org/10.1007/s10212-013-0205-2

Kraft, M. A., & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review, 47*(3), 49-63. https://doi.org/10.1016/j.econedurev.2015.04.001

Kulakow, S. (2020). Academic self-concept and achievement motivation among adolescent students in different learning environments: Does competence-support matter? *Learning and Motivation, 70*(1), 101-112. https://doi.org/10.1016/j.lmot.2020.101632

Kyriakides, L. (2005). Extending the comprehensive model of educational effectiveness by an empirical investigation. *School Effectiveness and School Improvement, 16*(2), 103-152. https://doi.org/10.1080/09243450500113936

Lei, H., Cui, Y., & Chiu, M. M. (2018). The relationship between teacher support and students' academic emotions: A meta-analysis. *Frontiers in Psychology, 8*(2), 22-88. https://doi.org/10.3389/fpsyg.2017.02288

Liu, W. C., & Wang, C. K. J. (2005). Academic self-concept: A cross-sectional study of grade and gender differences in a Singapore secondary school. *Asia Pacific Education Review, 6*(1), 20-27.

LoCasale-Crouch, J., Jamil, F., Pianta, R. C., Rudasill, K. M., & DeCoster, J. (2018). Observed quality and consistency of fifth graders' teacher–student interactions: Associations with feelings, engagement, and performance in school. *Sage Open, 8*(3), 1-11. https://doi.org/10.1177/2158244018794774

Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J., Abdelfattah, F., Nagengast, B., Möller , J., & Abu-Hilal, M. M. (2015). The internal/external frame of reference model of self-concept and achievement relations: Age-cohort and cross-cultural differences. *American Educational Research Journal, 52*(1), 168-202. https://doi.org/10.3102/0002831214549453

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*(2), 181-220. https://doi.org/10.1207/s15327906mbr3302_1

Marsh, H. W., & Martin, A. J. (2011). Academic self□concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59-77. https://doi/org /10.1348/000709910X503501

Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. G. (2017). Competence self-perceptions. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application*, (pp. 85-115). The Guilford Press.

McFarland, L., Murray, E., & Phillipson, S. (2016). Student-teacher relationships and student self-concept: Relations with teacher and student gender. *Australian Journal of Education, 60*(1), 5-25. https://doi.org/10.1177/0004944115626426

Niepel, C., Brunner, M., & Preckel, F. (2014). The longitudinal interplay of students' academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology, 106*(4), 1170-1191. https://doi.org/10.1037/a0036307

Pakarinen, E., Silinskas, G., Hamre, B. K., Metsäpelto, R. L., Lerkkanen, M. K., Poikkeus, A. M., & Nurmi, J. E. (2018). Cross-lagged associations between problem behaviors and teacher-student relationships in early adolescence. *The Journal of Early Adolescence, 38*(8), 1100-1141. https://doi.org/10.1177/0272431617714328

Parker, P. D., Marsh, H. W., Morin, A. J., Seaton, M., & Van Zanden, B. (2015). If one goes up the other must come down: Examining ipsative relationships between math and English self concept trajectories across high school. *British Journal of Educational Psychology, 85*(2), 172-191. https://doi/org /10.1111/bjep.12050

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120-141). Routledge/Taylor & Francis Group.

Pino-Pasternak, D., Whitebread, D., & Tolmie, A. (2010). A multidimensional analysis of parent-child interactions during academic tasks and their relationships with children's self-regulated learning. *Cognition and Instruction, 28*(3), 219-272. https://doi.org/10.1080/07370008.2010.490494

Pinxten, M., De Fraine, B., Van Damme, J., & D'Haenens, E. (2013). Student achievement and academic self-concept among secondary students in Flanders: Gender and changes over time. *Irish Educational Studies, 32*(2), 157-178. https://doi/org/10.1080/03323315.2012.749058

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology, 84*(1), 152-174. https://doi.org/10.1111/bjep.12028

Pitt, V., Powis, D., Levett-Jones, T., & Hunter, S. (2012). Factors influencing nursing students' academic and clinical performance and attrition: An integrative literature review. *Nurse Education Today, 32*(8), 903-913. https://doi/org /10.1016/j.nedt.2012.04.011

Pitt, V., Powis, D., Levett-Jones, T., & Hunter, S. (2014). Nursing students' personal qualities: A descriptive study. *Nurse Education Today, 34*(9), 1196-1200. https://doi/org / 10.1016/j.nedt.2014.05.004

Raccanello, D., & De Bernardi, B. (2013). A research study on the relationships between student achievement goals, perception of contextual achievement goals and academic performance in Italian and Mathematics. *BPA-Applied Psychology Bulletin (Bollettino di Psicologia Applicata), 61*(2), 268-281. http://hdl.handle.net/20.500.12424/2408110

Rady, H., Kabeer, S., & El-Nady, M. T. (2016). Relationship between academic self concept and students' performance among school age children. *American Journal of Nursing Science, 5*(6), 295-302.

https://doi/org /10.11648/j.ajns.20160506.19

Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher-child relationships. *Attachment & human development, 14*(3), 213-231. https://doi/org/10.1080/14616734.2012.672262

Sakiz, G. (2012). Perceived instructor affective support in relation to academic emotions and motivation in college. *Educational Psychology, 32*(1), 63-79. https://doi/org /10.1080/01443410.2011.625611

Schiefele, U., & Schaffner, E. (2015). Teacher interests, mastery goals, and self-efficacy as predictors of instructional practices and student motivation. *Contemporary Educational Psychology, 42*(3), 159-171. https://doi.org/10.1016/j.cedpsych.2015.06.005

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological bulletin, 143*(6), 565-600. https://doi/org /10.1037/bul0000098

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd Ed.). Taylor & Francis Group.

Simonsmeier, B. A., Peiffer, H., Flaig, M., & Schneider, M. (2020). Peer feedback improves students' academic aelf-concept in higher education. *Research in Higher Education, 61*(2), 706-724. https://doi.org/10.1007/s11162-020-09591-y

Skinner, E., Pitzer, J., & Brule, H. (2014). The role of emotion in engagement, coping, and the development of motivational resilience. In R. Pekrun & L. Linnenbrink-Garcia (Eds.) *International handbook of emotions in education*, (1st ed.), (pp. 331-347). Routledge Publication.

Trautwein, U., & Möller, J. (2016). Self-concept: Determinants and consequences of academic self-concept in school contexts. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the 21st century: Theory, research, and practice*, (pp. 187-214). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-28606-8_8

Veas, A., Castejón, J.-L., Miñano, P., & Gilar-Corbí, R. (2019). Actitudes en la adolescencia inicial y rendimiento académico: El rol mediacional del autoconcepto académico [Early adolescents' attitudes and academic achievement: The mediating role of academic self-concept]. *Revista de Psicodidáctica, 24*(1), 71-77. https://doi.org/10.1016/j.psicod.2018.11.001

Verschueren, K. (2015). Middle childhood teacher-child relationships: Insights from an attachment perspective and remaining challenges. *New directions for child and adolescent development, 148*(2), 77-91. https://doi/org /10.1002/cad.20097. PMID: 26086129

Verschueren, K., Doumen, S., & Buyse, E. (2012). Relationships with mother, teacher, and peers: Unique and joint effects on young children's self-concept. *Attachment & human development, 14*(3), 233-248. https://doi.org/10.1080/14616734.2012.672263

Verschueren, K., & Koomen, H. M. (2012). Teacher-child relationships from an attachment perspective. *Attachment & human development, 14*(3), 205-211. https://doi/org /10.1080/14616734.2012.672260

Vervoort, E., Doumen, S., & Verschueren, K. (2015). Children's appraisal of their relationship with the teacher: Preliminary evidence for construct validity. *European Journal of Developmental Psychology, 12*(2), 243-260. https://doi.org/10.1080/17405629.2014.989984

Westphal, A., Kretschmann, J., Gronostaj, A., & Vock, M. (2018). More enjoyment, less anxiety and boredom: How achievement emotions relate to academic self-concept and teachers' diagnostic skills. *Learning and Individual Differences, 62*(1), 108-117. https://doi.org/10.1016/j.lindif.2018.01.016

Wolff, F., Nagy, N., Helm, F., & Möller, J. (2018). Testing the internal/external frame of reference model of academic achievement and academic self-concept with open self-concept reports. *Learning and Instruction, 55*(3), 58-66. https://doi.org/10.1016/j.learninstruc.2017.09.006

Wubbels, T., & Brekelmans, M. (2005). Two decades of research on teacher-student relationships in class. *International Journal of Educational Research, 43*(1-2), 6-24. https://doi.org/10.1016/j.ijer.2006.03.003

Wubbels, T., & Levy, J. (1993). *Do you know what you look like? Interpersonal relationships in education*. Falmer Press/Taylor & Francis, Inc.

Journal of Language Horizons

Journal of Language Horizons, *Alzahra University*
Volume 6, Issue 1, Winter and Spring 2022 (Biannual – Serial No. 11)

The Role of Semantic Priming in Relative Clause Attachment Ambiguity Resolution in Persian

Research Article
pp. 247-261

Eskandar Samadi[1]
Mohammad Nabi Karimi*[2]
Esmat Babaii[3]

Received: 2020/06/19 Accepted: 2020/12/05

Abstract
The present study investigated the role of semantic priming in the processing of ambiguous sentences containing Relative Clauses (RCs) preceded by a complex Noun Phrase (NP) by Persian native speakers. To this end, in a self-paced reading task using E-prime software, 63 Persian native speakers read sentences containing ambiguous RCs in their L1, i.e. Persian. The type of semantic relationship in this study was the one between the RC and one of the NPs within the complex NP to find out whether priming one of the NPs through this semantic manipulation would affect Persian native speakers' attachment preferences. The results of the off-line post interpretive (RC attachment preferences) and on-line data (reading times) revealed that semantic priming affects participants' attachment preferences, which suggests that their parsing preferences are not guided purely by syntactic information. The findings are in line with constraint-based models of sentence parsing, which assume that, during parsing, multiple sources of information interact and each of them constrains the interpretation in a particular way. The results also support the predictions of the Spreading Activation Model and Lexical Priming Theory.

keywords: RC attachment preference, ambiguity resolution, semantics, priming, native Persian speakers

*Corresponding Author
[1] Ph.D. Candidate, Department of Foreign Languages, Kharazmi University, Tehran, Iran; Std_samadi@khu.ac.ir
[2] Associate Professor, Department of Foreign Languages, Kharazmi University, Tehran, Iran; Karimi_mn@khu.ac.ir
[3] Professor, Department of Foreign Languages, Kharazmi University, Tehran, Iran; Babaii@khu.ac.ir

DOI: 10.22051/LGHOR.2020.31707.1323

Introduction

According to Kecskes (2008, 2014), native speakers' ambiguity resolution preferences in a particular language reflect conventional ways of organizing thoughts. One type of ambiguity in this regard is relative clause (RC) attachment ambiguity which has attracted a growing surge of interest in the field of psycholinguistics since it examines the nature of human sentence processing mechanism (Marefat & Farzizadeh, 2018; Marefat & Samadi, 2015; Papadopoulou & Clahsen, 2003). The RC attachment ambiguity resolution paves the way for explaining the properties of the humans' parsing mechanisms and also determines different sorts of linguistic information used to resolve ambiguity. An example of RC attachment ambiguity is presented below:

(1) *Alex saw [the driver]$_{NP1}$ of [the manager]$_{NP2}$ [who was eating breakfast]$_{RC}$*

In this sentence, the ambiguous RC "*who was eating breakfast*" is preceded by a complex Noun Phrase (NP) "*the servant of the actress*", and both NPs within the complex NP have the potential to act as a host for the subsequent RC. The RC can be interpreted as modifying either the first NP (i.e., the driver), resulting in a High Attachment (HA), or the second NP (i.e., the manager), leading to a Low Attachment (LA).

There are different accounts which explain the NP1 and NP2 attachment preferences. The two mostly cited structure-based parsing models are Predicate Proximity and Recency. According to Papadopoulou (2005, p. 108) Predicate Proximity, alternatively named Early Closure, "requires new material to be attached as close as possible to the IP node" which is the first NP in the complex NP resulting in a HA. On the other hand, Recency, alternatively named Late Closure, "is assumed to be universal and forces the new material to be attached to the most recently processed phrase" (Papadopoulou, 2005, p. 108) which is the second NP in the complex NP resulting in a LA. The other parsing model which explains the NP1 and NP2 attachment preferences is based on multiple-constraints accounts (Green & Mitchell, 2006; MacDonald, 1994; Thornton et al., 1998; Traxler et al., 2000) which assume multiple sources of information, including semantic plausibility, subcategorization preferences, and discourse context, interact while each constrains the interpretation in a particular way.

The relevant literature provides evidence that cross-linguistic differences exist in the resolution of this kind of ambiguity (Papadopoulou & Clashen, 2003). High attachment (NP1) preferences have been reported in various languages including Arabic (Bidaoui et al., 2016), Dutch (Desmet et al., 2006), French (Colonna et al., 2000; Dekydtspotter et al., 2008), Greek (Papadopoulou & Clahsen, 2003), German (Hemforth et al., 2015; Hemforth et al., 2000) Russian (Iudina & Fedorova, 2009), Persian (Arabmofrad & Marefat, 2008; Shabani, 2016), and Spanish (Fernández, 2003). Also, low attachment (NP2) preferences have been reported in English (Bergmann et al., 2008), Japanese (Jun & Koike, 2008), Portuguese (Finger & Zimmer, 2000), Romanian, Swedish, and Norwegian (Ehrlich et al., 1999).

Previous research suggests that various factors may affect RC attachment preferences including prosody (Dekydtspotter et al., 2008; Fodor, 2002, Zahn & Scheepers, 2015), the kind of relativizing element (Hemforth, et al., 2000), animacy (Desmet & Declercq, 2006), the amount of exposure (Dekeyser, 2005; Caffarra et al., 2015), relative pronoun type (Delle Luche et al., 2006), availability of alternative structures (Mitchell et al., 2000), position and length of the RC (Hemforth et al., 2015), individual differences in working memory (WM) capacity (Kim &

Christianson, 2013; Marefat & Samadi, 2015; Traxler, 2007), proficiency (Miyao & Omaki, 2006), as well as semantics (Marefat & Samadi, 2015). However, previous studies have not examined whether semantic relationship between the RC and one of the NPs in a complex NP may affect parsers' attachment preferences, so the present study intends to fill some of the void in this regard.

Review of Literature

The Structure of Ambiguous RCs in Persian

Megerdoomian (2000, p. 5) stated that Persian complex NP is "the equivalent of genitive or possessive construction in English" in which two nouns are linked together by *'s* or the preposition *of*. In Persian, the *ezafe* morpheme *-e* (-ye after vowels) is an unwritten short vowel that connects the head of an NP to its following constituents. Therefore, a Persian complex NP is a sequence of simple NPs without any overt links or boundaries as shown in (2) below.

(2). *dust bæraadær Ali*
friend brother Ali
"A friend of Ali's brother"

The actual pronunciation is provided in (3).

(3). *dust-e bæraadær-e Ali*

Like English, Persian RCs are post-nominal which provide some information about a nominal element or "head" in the main clause. However, unlike English, in Persian, there is no relative pronoun and the RCs are typically introduced by the complementizer *ke* which is the sole complementizer in Persian (Taleghani, 2008). Unlike its English equivalent, the Persian complementizer is obligatory; and it is always used at the beginning of all RCs regardless of animacy, function, gender or the number of the head nouns modified by the RC (Rahmany, Marefat, & Kid, 2011; Taghvaipour, 2004). Persian complementizer *ke* contains the semantic meaning of 'who', 'whom', 'which', 'whose' and 'that' in itself. Additionally, in Persian two NPs in the complex NP are linked together by *ezafe* morpheme *-e* and the ambiguous RC could be attached to either of these NPs in the complex NP. Moreover, the ambiguous RC always precedes the object marker *ra* in Persian. According to the Predicate Proximity account, Persian is a language that favors high attachment since Persian has a relatively free-word order which allows the verb to be distant from its complements (Faghiri et al., 2014). Therefore, the RC is attached to the first NP. An example of Persian ambiguous RC followed by a complex NP is presented in the following example:

(4).

رضا وکیل دکتر که داشت ناهار میخورد را دید.

Reza vækil e doctor ke dasht nahar mikhord ra did.
Reza lawyer ezafe morpheme doctor who was lunch eat object marker saw.
'Reza saw the lawyer of the doctor who was eating lunch.'

In this sentence, the ambiguous RC "*ke dasht nahar mikhord"* could be attached to either of NPs (i.e., *vækil or doctor) in the complex NP* (i.e., *vækile doctor)* which are linked together by *ezafe* morpheme *-e.* Therefore, this sentence could be interpreted as either *"vækil dasht nahar mikhord" (i.e., the lawyer was eating lunch)* resulting in a high attachment (NP1) or *"doctor dasht nahar mikhord" (i.e., the doctor was eating lunch)* leading to a low attachment (NP2).

Theoretical Backgrounds for the Present Study

The purpose of this study is to determine whether semantic priming can affect Persian native speakers' RC attachment preferences. The type of semantic manipulation in the present study is the one between the RC and one of the NPs in the complex NP. It is possible for an RC to be semantically biased toward NP1 or NP2. Consider the following sentence as an example:

5) *Alex saw the [the patient]$_{NP1}$ of [the nurse]$_{NP2}$ [who was very pale]$_{RC.}$*

In this sentence, there is a semantic relationship between NP1 and RC which may bias NP1 attachment preferences on the basis of the Spreading Activation Model (Colins & Loftous, 1975; Traxler et al., 2000). The Spreading Activation is a theory of how the human brain reiterates through a network of connected ideas to recover distinct information. This theory offers a range of ideas and concepts within our memory as cognitive units, each comprising a node and its related characteristics or elements, all connected by edges (Anderson, 1983). In a sort of web diagram, the spreading activation network could be explained schematically; shorter lines between two nodes imply that these ideas are more closely related to each other and they would naturally be associated more quickly to the original concept. From memory psychology aspects of Spreading Activation Model, on the basis of their personal experience, individuals construct their knowledge of the world that form their network of ideas which is considered as their knowledge of the world (Colins & Loftous, 1975; Traxler et al., 2000). Snowden (2015) stated that the spreading activation theory of semantic processing represents a less firmly structured revision of Quillian's (1966) network model. This model integrates the idea of semantic distance, based on which highly related concepts are located closer together compared to unrelated ones. Therefore, based on this model, when a word is activated, other words which are semantically related to it also become activated. Accordingly, when NP1, *"the patient",* in case of sentence (5) is activated, other words that are semantically related to it (i.e., *being pale)*, also become activated; and these related words boost the activation of each other and as a result, the first NP (i.e., *the patient*) remains more accessible compared to NP2 "*the nurse"*. When the reader encounters the ambiguous RC (*who was very pale*"), s/he is expected to (if this theory is operative) attach the ambiguous RC to the more available NP, which is the first NP "the patient". Moreover, based on Lexical Priming Theory (Hoey, 2005) which states that large numbers of prefabricated elements are stored as units in the brain and primed by repeated encounters; it is more plausible for a patient to be pale than for a nurse, and this semantic relationship may bias NP1 attachment preference.

Similarly, the semantic relationship between NP2 and RC may bias NP2 attachment preferences.

6) *Alex saw the [the patient]$_{NP1}$ of [the nurse]$_{NP2}$ [who was filling the syringe]$_{RC.}$*

In sentence (6), there is a semantic relationship between NP2 and RC. Again, based on Lexical Priming Theory (Hoey, 2005), it is more reasonable for a nurse to fill a syringe than for a patient, and this semantic relationship may bias NP2 attachment preference. Moreover, based on the Spreading Activation Model, in this sentence, NP2 is more accessible compared to NP1, which makes it a more probable host for the following ambiguous RC.

Thus, there are hints in the literature indicating that when the RC is semantically biased towards one of the NPs in the complex NP, that NP becomes

more salient, causing the RC to be attached to it. And if the semantics does not affect participants' parsing preferences it could be concluded that their attachment preference is guided solely by syntactic information.

Against the background presented, this study aims to investigate the role of semantic
priming in RC attachment preferences of Persian native speakers. More specifically, the study aims to answer the following question:

Does priming one of the NPs within the complex NP through semantically associating it with the following RC influence RC attachment preferences of Persian native speakers?

Method

Participants

Fifty-six Persian speaking monolinguals (mean age 22, range 16–42, 23 females, 33 males) took part in this experiment. Three participants were excluded because they had not completed the task or did not fulfill the criterion of 85% comprehension accuracy of filler sentences (see the Results section for details). Therefore, for data analysis, the data from 53 participants were used. Twenty-five participants provided data on Version 1 of the Paraphrase Decision Task (PDT) and twenty-eight participants provided data on Version 2 of the task.

Materials

Paraphrase Decision Task (PDT). Two versions of PDT were used. Each version consisted of 70 sentences, including 28 test sentences and 42 fillers. In PDT, the stimuli were presented in a self-paced, chunk-by-chunk, noncumulative fashion. The PDT was implemented using E-Prime software.

Test Sentences. The experimental sentences used in this study were all structurally ambiguous sentences in Persian (the native language of the participants) containing a main clause and an ambiguous RC which could refer to either of the two preceding NPs that were linked together by the *ezafe* morpheme *-e in Persian*. In all test sentences, the ambiguous RC was introduced by the complementizer *ke*. Based on the semantic relationship between the RC and either of the two NPs in the complex NP, test sentences were categorized into three conditions:
NP1-biased: in which the RC was semantically related to NP1 (sentence 7 below);
NP2-biased: in which RC was semantically related to NP2 (sentence 8 below);
Unbiased: in which the RC could be linked to both NP1 and NP2 (sentence 9 below).

(7)

نگار معلم دانش آموز که ریاضی درس میداد را شناخت.

Negar moœlem e danesh-amouz ke riazi dars midad ra šenakht.

Negar teacher ezafe morpheme student who mathematics was teaching object marker recognized.

'Negar recognized the teacher of the student who was teaching mathematics.'

(8)

نگار معلم دانش آموز که در حیاط بازی میکرد را شناخت.

Negar moœlem e danesh-amouz ke dær hœyat bazi mikard ra šenakht.

Negar teacher ezafe morpheme student who in yard was playing object marker recognized.
'Negar recognized the teacher of the student who was playing in the yard.'
(9)

نگار معلم دانش آموز که به خانه میرفت را شناخت.
Negar moœlem e danesh-amouz ke be khaneh miræft ra šenakht.
Negar teacher ezafe morpheme student who to home was going object marker recognized.
'Negar knew the teacher of the student who was going home.'

The rationale for including Unbiased items in the study was to determine participants' general parsing preferences since in these sentences, there is no semantic relationship between the RC and either of NPs in the complex NP. Test sentences (7), (8), and (9) are considered as one set. In a norming study, twenty sets were developed into a questionnaire. It was distributed among 25 participants from a similar population as in the main study to evaluate the relationship set by the researchers between the RC and either of the NPs. None of these participants took part in the main study. Moreover, six experts in the field checked the content validity of the survey. As each set contains three sentences, there were 60 items altogether in the questionnaire. In each item, the RC was written in bold and the two NPs were underlined. Each item was followed by three choices: NP1, NP2, and both. The test takers were asked to decide which NP was related to the RC written in bold. An example is provided below:
(10)

نگار معلم دانش آموز ***که ریاضی درس میداد*** *را شناخت.*
معلم ☐ دانش‌آموز ☐ هردو ☐

Negar recognized the teacher of the student who was teaching mathematics.
the teacher ☐ the student ☐ both ☐

The test sentences would be included in the set of sentences for the main experiment if the corresponding choice had been selected by 90% of the respondents. More specifically, for the test item to be qualified as an NP1-biased RC, the first choice (i.e., NP1), should have been selected by 90% of the testees. Likewise, to be qualified as an NP2-biased RC, the second choice (i.e., NP2), needed to be selected by 90% of the testees. Finally, to be qualified as an Unbiased RC, the third choice (i.e., both), had to be selected by 90% of the testees.

Fourteen sets which met the criterion were used in the main experiment. In order to use all the fourteen sets, test sentences were presented in two versions, *Version 1* and *Version 2,* to reduce test fatigue and to avoid participants' test awareness. Each version included 14 NP-biased sentences (i.e., seven NP1 and seven NP2 RC-biased sentences). Moreover, for each NP-biased item, there was an Unbiased item. Accordingly, there were fourteen Unbiased RCs in each version. If the NP1-biased-RC item of a set was in *Version 1* of the main test, in *Version 2* it was replaced with the NP2-biased-RC item and vice versa. The Unbiased-RC item of each set was common in the two versions.

Test sentences were divided into four fragments, sentence (7) is repeated below in (11) as an example:
(11)

نگار/ معلم دانش آموز /که ریاضی درس میداد/ را شناخت.
4 3 2 1

Negar/ moœlem e danesh-amouz/ ke riazi dærs midad/ ra šenakht.
1 2 3 4
'Negar knew the teacher of the student who was teaching mathematics.'

This fragmentation is motivated by a need to prevent participants from assigning prosodic structure to the sentences presented and also to encourage natural reading of sentences. In these items, as shown in sentence 11, the third and the fourth fragments (i.e., the RC, the object marker *ra*, and the following verb) were the critical regions. The third fragment (i.e., the RC), is the part in which semantics is manipulated. The RC is followed by a verb in order to be able to observe possible spill-over effects, so the fourth segment (i.e., the verb) is considered as the post-critical region. As mentioned before the NP chosen by the participants in the Unbiased condition reveals their general parsing preference since there is no semantic association between the ambiguous RC and either of NPs in the complex NP. The prediction would be that if participants prefer a high (NP1) attachment, then their reading times for critical regions in sentences in which the RC is semantically associated with NP2 would take longer compared to sentences in which the RC is semantically associated with NP1. Conversely, if they favor attaching the RC to NP2 (i.e., low attachment) then, their reading time for sentences in which the RC is semantically associated with NP1 would take longer. Each test sentence was followed by a paraphrase which referred to either NP1 or NP2. Both paraphrases were accurate, half of the paraphrased sentences were disambiguated towards NP1 and the other half were disambiguated towards NP2. An example of a test sentence, together with its paraphrase, is presented below.

(12)

فرنوش آرایشگر هنرپیشه که فیلم نامه تمرین میکرد را تحسین کرد.

آرایشگر فیلمنامه تمرین میکرد.

Farnooš arayešgar e honærpišeh ke filmnameh tæmrin mikærd ra tæhsin kærd.
Farnooš barber ezafe morpheme actress who script was practicing object marker adored.
'Farnooš adored the barber of the actress who was practicing the script.
Arayešgare filmnameh tæmrin mikærd.
The barber script was practicing
'The barber was practicing the script.'

As observed, this sentence is disambiguated towards NP1 (i.e., high attachment). As another example, consider sentence (13) below in which the paraphrased sentence is disambiguated towards NP2 (i.e., low attachment).

(13)

فرنوش آرایشگر هنرپیشه که فیلم نامه تمرین میکرد را تحسین کرد.

هنرپیشه داشت فیلمنامه تمرین میکرد.

Farnooš arayešgar e honærpišeh ke filmnameh tæmrin mikærd ra tæhsin kærd.
Farnooš barber ezafe morpheme actress who script was practicing object marker adored.
'Farnooš adored the barber of the actress who was practicing the script.

Honærpišeh filmnameh tæmrin mikærd.
The actress script was practicing
'The actress was practicing the script.'

In (12), if the participants press *true* key, it means that semantic association does not affect their parsing preference and if they press *false* key, this shows that semantic association affects their parsing preference. On the other hand, in (13), if the participants press *true* key, this means that semantic association affects their parsing preference, and if they press *false* key, this shows that semantic association does not affect their parsing preference

Filler Sentences. Forty-two filler sentences were developed for this study. Fillers had various grammatical structures including unambiguous RCs and were matched with experimental sentences in length. The fillers were used for distracting the participants from becoming aware of the purpose of the study. Similar to experimental sentences, each filler was followed by a paraphrased sentence which was correct in half of the items and wrong in the other half. An example of a filler item with its paraphrase is provided below:

(14)

هفته گذشته منشی اداره با اتوبوس به اصفهان رفت.

منشی با قطار به اصفهان رفت.

Hæfteh gozæshšteh monšie edæreh ba otobus be Esfæhan raft.
Last week the secretary the office by bus to Esfahan went.
'last week the secretary of the office went to Esfahan by bus.'
Monši ba ghætar be Esfæhan ræft.
'The secretary went to Esfahan by train.'

In all sentences, including test sentences and filler items, participants were required to determine whether the paraphrased sentence was correct or not. The paraphrased sentences served two purposes. First, to find out which NP was selected by participants as the host of the following RC; and second, to check whether the participants paid enough attention to the content of the test or not. Considering filler sentences, the paraphrased sentences only served the second purpose because their answers to paraphrases following the fillers could be checked only for accuracy. The participants with less than 85% accuracy in responses to fillers were excluded from the study. The participants' answers to paraphrases following the test sentences just indicated their attachment preferences, and thus could not be checked for accuracy.

Each version of the test included 70 items including 28 test sentences (including seven NP1-biased, seven NP2-biased, and fourteen Unbiased items) and 42 fillers. In order to prevent test fatigue, each version was divided into two halves and was presented to the participants in two sessions. Each participant took the first half including 35 items then after a short break s/he took the second half.

Practice Test. Before taking the main test, each participant took five warm-up sentences in order to get familiar with how to proceed with the experiment. These sentences were the same across the two versions. Like experimental sentences, the participants were asked to determine whether the paraphrases following each sentence were correct or not. The participants enjoyed the liberty to ask questions regarding the sentences, software, etc. Having mastered the procedure, the participants took the main test.

Procedure

The participants were tested individually by an 18-inch laptop. First, the five-item practice test was administered to familiarize them with the task. Afterwards, the participants were given the main test containing 70 items. Using 22-point Times New Roman font, the stimuli were presented in a noncumulative, chunk-by-chunk, self-paced method. The participants received each sentence in four segments as indicated by the slashes in (11). The participants were also instructed that by pressing the space button on the keyboard, a segment would appear on the laptop screen. Each segment stayed in the middle of the laptop screen until participants' next key-press. After each key-press, the segment that the participants read disappeared and the then the following section showed up on the screen. This process continued until the last segment of the sentence appeared on the screen. When the participants finished reading each sentence, a true/false statement appeared on the screen. The participants were asked to determine whether that statement was correct or incorrect by pressing "Right arrow" or "Left arrow" buttons on the keyboard, respectively. If the 'Right arrow' was chosen in response to an HA paraphrase, it was interpreted as a sign of HA preference. If the 'Left arrow' was chosen, the preference was considered as LA. If the 'Right arrow' was the response to an LA paraphrase, that would mean a LA preference. Finally, if the 'Left arrow' was chosen in response a LA paraphrase, it would be an indication of a HA preference. No feedback was given regarding participants' responses. Moreover, participants' reading times (RTs) for each segment in each sentence and their answers to true/false paraphrased statements were recorded automatically by the software in milliseconds.

Results

Prior to data analyses, participants' answers to the true/false statements following the fillers were checked to ensure that they had paid enough attention to the content of the test and read sentences carefully. Those participants whose accuracy scores were lower than 85% (three participants) were excluded from the analyses. On average, the participants answered 92.31 % of fillers correctly.

Results for Post-Interpretive Offline Data

In order to find out whether semantics affects participants' attachment preferences, the percentages of their NP1 and NP2 choices across the three conditions were calculated. Participants' answers to sentences in the Unbiased condition reveals their parsing preferences in general. In the Unbiased condition, as Table 1 below shows, 87.06% of the responses referred to NP1 while only 12.94% of the responses referred to NP2 which clearly shows that they have a HA preference. Similarly, in the NP1-biased condition, 97.3% of responses referred to NP1, but just 2.69% of the responses referred to NP2. Contrary to these two conditions, in the NP2-biased condition, only 13.2% of the responses referred to NP1 and 87.87% of responses referred to NP2. These results clearly indicate that semantics affects participants' attachment preferences.

Table 1
Percentage of NP1 and NP2 Selection Across the Three Biased Conditions

Condition	Antecedent	
	NP1	NP2
Unbiased	87.06%	12.94%
NP1-biased	97.30%	2.69%
NP2-biased	13.20%	87.87%

Results for On-line Data (Reaction Times)

In the next step, participants' reaction times (RTs) for reading the critical region (i.e., region 3) and the post-critical region (i.e., region 4) in the three conditions were compared. Before analyzing the data, participants' RTs for each region were divided by the number of syllables in that region in order to normalize the RCs and the following regions for their differences in length (Carreiras & Clifton, 1999). Moreover, the distribution of the time was detected for outliers for each participant in each of the three conditions and was substituted by the mean RT for that participant in the condition where the outlier was located.

Results for RCs. Table 2 below presents the descriptive statistics for participants' mean RTs for reading ambiguous RCs across the three conditions (i.e., Unbiased, NP1-biased, NP2-biased).

Table 2
Mean RTs for the RC Across the Three Conditions

	Mean	Std. Deviation	N
Unbiased	380.4776	46.27222	53
NP1-biased	337.2915	54.41719	53
NP2-biased	819.9768	176.30949	53

Mauchly's Test of Sphericity revealed that the assumption of sphericity had been violated, $\chi^2(2) = 72.89$, $p = 0.000$. Therefore, the degrees of freedom were corrected using Greenhouse-Geisser correction estimates ($\varepsilon = .56$). Repeated measures ANOVA determined that mean RTs for the ambiguous RCs (i.e., *region 3*) differed statistically significantly between three conditions (i.e., *Unbiased, NP1-biased*, and *NP2-biased*) [$F(1.136, 59.07) = 322.87$, $p = .000$, $\eta_p^2 = .861$]. Post hoc tests using the Bonferroni correction revealed that there were significant differences among RTs in the three conditions. The participants produced shorter RTs for sentences in which, through semantic manipulation, RC was semantically associated with NP1 ($M = 337.29$, $SD = 54.41$), but very longer RTs to sentences in which the RC was semantically associated with NP2 ($M = 819.97$, $SD = 176.3$). Moreover, there was also a slight but significant difference between the Unbiased condition ($M = 380.47$, $SD = 46.27$) and NP1-biased condition in which RC was biased towards NP1. There was also a significant difference between the Unbiased condition ($M = 380.47$, $SD = 46.27$) and NP2-biased condition ($M = 819.97$, $SD = 176.3$) in which RC was biased towards NP2. Although participants had a HA preference, the RTs for reading the NP1-biased sentences were significantly shorter than the Unbiased condition meaning that semantic manipulation helped them read and comprehend these sentences faster.

Results for Spillover Region. Table 3 below presents the descriptive statistics for participants' mean RTs for reading spillover region across the three conditions (i.e., Unbiased, NP1-biased, NP2-biased).

Table 3
Mean RTs for the Spillover Region Across the Three Conditions

	Mean	Std. Deviation	N
Unbiased	1380.7822	130.97646	53
NP1-biased	1099.2471	219.31002	53
NP2-biased	2075.4410	634.22941	53

In the final step, the spillover effect was examined among the three conditions. Mauchly's test revealed that the assumption of sphericity had also been violated, in case of spillover region, $\chi 2(2) = 73.167$, $p = .000$. Again, a repeated measures ANOVA with a Greenhouse-Geisser correction determined that mean RTs for the post-critical region (i.e., *region 4*) differed statistically significantly among the three mentioned conditions $F(1.135, 59.03) = 75.88$, $p = .000$, $\eta_p^2 = .593$). Post hoc tests using the Bonferroni correction indicated that there were significant differences between the RTs in the three conditions. The participants produced shorter reaction times for sentences in which due to a semantic manipulation, RC was semantically associated with NP1 ($M = 1099.24$, $SD = 219.31$), but longer RTs to sentences in which the RC was semantically associated with NP2 ($M = 2075.44$, $SD = 634.22$). Additionally, there was also a significant difference between RTs in the Unbiased condition ($M = 1380.78$, $SD = 130.97$) and the NP1-biased condition ($M = 1099.24$, $SD = 219.31$). Moreover, there was a significant difference between RTs in the Unbiased condition ($M = 1380.78$, $SD = 130.97$) and the NP2-biased condition ($M = 2075.44$, $SD = 634.22$).

Discussion

The present study examined whether RC ambiguity resolution by Persian native speakers is sensitive to semantic priming. The type of semantic manipulation investigated in this study was the one between the RC and one of the NPs in the complex NP. The results suggest that semantic priming between the RC and one of the NPs in the complex NP has clearly affected participants' attachment preference. The NP chosen by the Persian native speakers in the Unbiased condition (i.e., NP1) revealed that they had a HA preference in general. In the Unbiased condition, participants' attachment preference seems to be guided by the principles of Predicate Proximity which favors the attachment of RC to NP1. These findings are in line with previous studies (Arabmofrad & Marefat, 2008; Shabani, 2016) which have showed that in Persian language where adjuncts can occur between the verbs and their complements, principles of Predicate Proximity are more operative compared to Late Closure. The results also showed that when there was a semantic relationship between the RC and NP1, the participants chose NP1 as the host of the following RC and when there was a semantic relationship between the RC and NP2, the participants chose NP2 as the host of the following RC. The findings are consistent with constraint-based models (Green & Mitchell, 2006; MacDonald 1994; Thornton et al., 1998; Traxler et al., 2000) which assume that several sources of information including discourse context and semantic plausibility affect processing of ambiguous sentences, while each of them constrains the interpretation in a particular way during the processing. Therefore, in addition to the phrase structure information (Dussias, 2003; Papadopoulou & Clahsen, 2003) which has proven to affect relative clause attachment ambiguity resolution, lexical and discourse information including semantic information, may exert an influence on participants' sentence processing.

The findings of this study are also in line with the Spreading Activation Model's predictions (Colins & Loftous, 1975; Traxler et al., 2000) which posits that when a word is activated other words which are semantically related to it also become activated and these words (the NPs and the biased RCs) boost the activation of each other and make the NP to which the RC is biased towards more accessible one and make the parser attach the RC to it. Moreover, the findings are in line with Lexical Priming Theory (Hoey, 2005) which states that large numbers of prefabricated elements are stored as units in the brain and primed by repeated encounters.

Although the general picture is very clear, it would be interesting to see that participants' attachment preferences for RCs in three different conditions figured differently in the RTs measures we inspected. Processing cost for critical and post-critical regions in the NP2-biased condition was enhanced because participants' initial attachment (i.e., HA) had to be revised since the RC was semantically biased towards a non-preferred attachment (i.e., LA); therefore, they had to refixate their initial NP1 attachment as the host of the following RC which leads to longer RTs. In the Unbiased condition, the participants, based on Predicate Proximity principle, attached the ambiguous RC to the first NP. Since no reanalysis was necessary, they did not recheck NP1 attachment for RCs as it was the preferred one; accordingly, they produced shorter RTs. Moreover, in the NP1-biased condition, the participants produced shorter RTs even compared to the Unbiased condition. This means that although they had a HA preference in general, semantic manipulation helped them process and comprehend NP1-biased sentences faster.

The present study provides room for further research. The results of the study could be extended by replicating it in other languages in order to cross-validate these findings considering the role of semantics. This study could be replicated with L1-Persian learners of L2 English or even L1-English learners of L2 Persian with different levels of proficiency to assess whether their attachment preference is affected by their L2. Omaki and Ariji (2004) posit that among various issues addressed in L2 research, the question of how an L2 learner's grammar progresses from one stage to another in the course of time has received little attention. They proposed that investigating the ways L2 learners analyze and parse the target language input and comparing the results with those of native speakers of that language would provide researchers with some insights into how L2 learners restructure target language input in non-native-like ways. Moreover, there could be a more complete and more informative research by altering the dependent and independent variables considering different aspects of structural ambiguity and a broader range of participants. Future research could also focus on the type of tasks, for example, eye-tracking and event-related potential to triangulate the results. This study could be accompanied by special tests which measure Working Memory Capacity (WMC) of readers. As a result, researchers could investigate whether WMC has any influence on the type of RC attachment preferences studied in this study; such a consideration can hopefully supplement the available evidence on processing to build a more complete picture.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261-295. https://doi.org/10.1016/S0022-5371(83)90201-3

Arabmofrad, A. & Marefat, H. (2008). RC attachment ambiguity resolution in Persian. *Iranian Journal of Applied Linguistics, 12*(1), 29-49. http://ijal.khu.ac.ir/article-1-71-en.html

Bergmann, A., Armstrong, M., & Maday, K. (2008). Relative clause attachment in English and Spanish: A production study. In *Proceedings of Speech Prosody* (pp. 505-518). http://isle.illinois.edu/sprosig/sp2008/papers/id147.pdf

Bidaoui, A., Foote, R., & Abunasser, M. (2016). Relative clause attachment in native and L2 Arabic. *International Journal of Arabic Linguistics*, *2*(2), 75-95. https://revues.imist.ma/index.php/IJAL/article/view/11564

Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: An empirical review. *Neuroscience & Biobehavioral Reviews*, *51*(1), 31-47. https://doi.org/10.1016/j.neubiorev.2015.01.010

Carreiras, M., & Clifton, C. (1999). Another word on parsing relative clauses: Eye tracking evidence from Spanish and English. *Memory & cognition*, *27*(5), 826-833. https://doi.org/10.3758/bf03198535

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428. https://doi.org/10.1037/0033295X.82.6.407

Colins, A. M., & Loftus, E. F. (1999). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428. https://eric.ed.gov/?id=EJ135579

Colonna, S., Pynte, J., & Mitchell, D. C. (2000, March 30-April 1). Relative clause attachment in French: The role of constituent length. *13th CUNY conference on human sentence processing,* La Jolla.

Dekeyser, R. M. (2005). What makes learning second language grammar difficult? A review of issues. *Language learning*, *55*(1), 1-25. https://doi.org/10.1111/j.0023-8333.2005.00294.x

Dekydtspotter, L., Donaldson, B., Edmonds, A. C., Fultz, A. L., & Petrush, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners. *Studies in Second Language Acquisition*, *30*(4), 453-480. https://doi.org/10.1017/S0272263108080728

Delle Luche, C., Van Gompel, R. P., Gayraud, F., & Martinie, B. (2006). Effect of relative pronoun type on relative clause attachment. In R. Artstein & M. Poesio (Eds.), *Ambiguity in Anaphora workshop proceedings* (p. 23-30). Malaga.

Desmet, T., DeBaecke, C., Drieghe, D., Brysbaert, M., & Vonk, W. (2006). Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, *21*(4), 453-485. https://doi.org/10.1080/01690960400023485

Desmet, T., & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language*, *54*(4), 610-632. https://doi.org/10.1016/j.jml.2005.12.007

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, *25*(4), 529-557. https://doi.org/10.1017/S0272263103000238

Ehrlich, K., Fernandez, E., Fodor, J. D., Stenshoel, E., & Vinereanu, M. (1999, March 18-20). *Low attachment of RCs: New data from Swedish, Norwegian, and Romanian*

[Poster presentation]. The 12th annual CUNY Conference on Human Sentence Processing, New York.

Faghiri, P., Samvelian, P., & Hemforth, B. (2014). Accessability and word order: The case of ditransitive constructions in Persian. In S. Müller (Ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar* (pp. 217-237). CSLI Publications.

Fernández, E. M. (2003). *Bilingual sentence processing: RC attachment in English and Spanish*. John Benjamins Publishing.

Finger, I., & Zimmer, M. (2000). Relative clause attachment preference in Brazilian Portuguese. Unpublished manuscript.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In B. Bell & I. Marlien (Eds.), *Speech Prosody: Proceedings, Aix-en-Provence, France* (pp. 83-88). Aboratoire Parole et Language – SPROSIG.

Green, M. J., & Mitchell, D. C. (2006). Absence of real evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *55*(1), 1-17. https://doi.org/10.1016/j.jml.2006.03.003

Hemforth, B., Fernández, S., Clifton Jr, C., Frazier, L., Konieczny, L., & Walter, M. (2015). Relative clause attachment in German, English, Spanish and French: Effects of position and length. *Lingua*, *166*, 43–64. https://doi.org/10.1016/j.lingua.2015.08.010

Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In M. W. Crocker, M. Pickering, & C. Clifton, Jr. (Eds.), *Architectures and mechanisms for language processing* (pp. 259-81). Cambridge University Press.

Hoey, M. (2005). *Lexical Priming: A new theory of words and language*. Routledge.

Iudina, M. V., Fedorova, O. V. (2009). Syntactic ambiguity resolution: Priming and self-priming effects. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue 2009*", *8*(15), 554-559.

Jun, S. A., & Koike, C. (2008). Default prosody and relative clause attachment in Japanese. *Japanese-Korean Linguistics*, *13*(1), 41-53. https://linguistics.ucla.edu/people/jun/JunKoike-final.pdf

Kecskes, I. (2008). The effect of the second language on the first language. *Babylonia*, *2*(2), 31-34. https://www.albany.edu/faculty/ikecskes/files/babyloniaproofkecskes.pdf

Kecskes, I. (2014). *Intercultural pragmatics*. Oxford University Press.

Kim, J. H., & Christianson, K. (2013). Sentence complexity and working memory effects in ambiguity resolution. *Journal of Psycholinguistic Research*, *42*(5), 393-411. https://doi.org/10.1007/s10936-012-9224-4

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157-201. https://doi.org/10.1080/01690969408402115

Marefat, H., & Farzizadeh, B. (2018). Relative clause ambiguity resolution in L1 and L2: Are processing strategies transferred? *Iranian Journal of Applied Linguistics*, *21*(1), 125-161. http://ijal.khu.ac.ir/article-1-2855-en.html

Marefat, H., & Samadi, E., (2016). Semantic priming effect on relative clause attachment ambiguity resolution in L2. *Applied Research on English Language, 4*(2), 79-94. https://doi.org/10.22108/are.2015.15504

Megerdoomian, K. (2000). *Persian computational morphology: A unification-based approach*. Computing Research Laboratory, New Mexico State University.

Mitchell, D. C., Brysbaert, M., Grondelaers, S., & Swanepoel, P. (2000). Modifier attachment in Dutch: Testing aspects of construal theory. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 493-516). Elsevier Science Publishers.

https://doi.org/10.1016/B978-008043642-5/50023-1

Miyao, M., & Omaki, A. (2006). No ambiguity about it: Korean learners of Japanese have a clear attachment preference. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.) *A supplement to the proceedings of the 30th Annual Boston University Conference on Language Development*.
http://www.bu.edu/bucld/files/2011/05/30-MiyaoBUCLD2005.pdf

Omaki, A., & Ariji, K. (2004). Testing and attesting the use of structural information in L2 sentence processing. In L. Dekydtspotter, R. A. Sprouse & A. Liljestrand (Eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference* (pp. 205-218). Cascadilla Proceedings.

Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Second Language Research*, *21*(2), 98-120.
https://doi.org/10.1191/0267658305sr254oa

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *25*(4), 501-528.
https://doi.org/10.1017/S0272263103000214

Quillian, M. R. (1966). *Semantic memory*. Bolt, Beranak & Newman.

Shabani, K. (2016, March). RC-attachment ambiguity resolution in Persian sentences [Paper presentation]. *The Second International Conference on Behavioral Science and Social Studies*, Istanbul, Turkey.

Snowden, J. S. (2015). Semantic memory. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (pp. 572-578). Elsevier.

Taghvaipour, M. (2004). An HPSG analysis of Persian relative clauses. In S. Muller (Ed.), *Proceedings of the HPSG-2004 conference, center for computational linguistics, Katholieke Universiteit Leuven,* (pp. 274-293). CSLI Publications.

Taleghani, A. H. (2008). *Modality, aspect and negation in Persian* (Vol. 128). John Benjamins Publishing.

Thornton, R., Gil, M., & MacDonald, M. (1998): Accounting for cross-linguistic variation: A constraint-based perspective. In Hillert, D. (Ed.), *Sentence processing: A crosslinguistic perspective. Syntax and semantics* (volume 31) (pp. 211-223). Academic Press.

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modeling analysis. *Memory & Cognition*, *35*(5), 1107-1121.
https://doi.org/10.3758/BF03193482

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intra-lexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research*, *29*(6), 581-595.
https://doi.org/10.1023/A:1026416225168

Traxler, M. J., Pickering, M. J., Clifton Jr, C., & van Gompel, R. (2000). Architectures and mechanisms for sentence processing: Is syntactic parsing a form of lexical ambiguity resolution? In *Cross-linguistic perspectives on language processing* (pp. 149-174). Springer.

Zahn, D., & Scheepers, C. (2015). *Overt prosody and plausibility as cues to relative-clause attachment in English spoken sentences*. Peerj Preprints.
http://doi.org/10.7287/peerj.preprints.1210v1

Abstracts of Papers in Persian

مجله افق‌های زبان
LGHOR
دانشگاه الزهرا

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

بررسی ارتقای متنی داده و تولید برون‌داد در رابطه با توجه به بندهای موصولی انگلیسی: مورد فراگیران زبان انگلیسی ایرانی

مقاله پژوهشی
صفحات ۲۶-۷

سمیه صادقی[۱]
پرویز مفتون[*۲]
مسعود یزدانی مقدم[۳]

تاریخ دریافت: ۱۳۹۹/۰۴/۲۳ تاریخ پذیرش: ۱۳۹۹/۱۱/۱۸

چکیده

طیِ دهه‌های اخیر، محققین فراگیری زبان دوم، توجه را یک فرایند شناختی مهم میانجیگرِ یادگیری زبان دوم در نظرگرفته‌اند. مفهوم توجه کردن به مثابه توجه آگاهانه به صورت، به ویژه در این حوزه، مفهوم مهمی شده‌است. پژوهش حاضر، به دنبال آن است که آیا دو مداخله آموزشی، به نام ارتقای متنی داده و تولید برون‌داد زبان‌آموزان، به تنهایی و به صورتی ترکیبی، توجه کردن زبان‌آموز و یادگیری بندهای موصولی انگلیسی را ارتقاء می‌دهد. مطالعه بر نمونهٔ ۱۱۳ نفری از دانشجویان سال اول رشته زبان و ادبیات انگلیسی انجام شد. شرکت‌کنندگان براساس نوع برخوردهایی که دریافت کردند به سه گروه آزمایشی و گروه کنترل دسته‌بندی شدند: ارتقای متنی داده و تولید برون‌داد زبان‌آموزان، ارتقای متنی داده و نبود تولید برون‌داد زبان‌آموزان، نبود ارتقای متنی داده و تولید برون‌داد زبان‌آموزان، نبود ارتقای متنی داده و نبود تولید برون‌داد زبان‌آموزان. پیش از انجام آزمایش، همه شرکت‌کنندگان آزمون بندهای موصولی انگلیسی به عنوان پیش‌آزمون را انجام دادند. سپس از شرکت‌کنندگان خواسته شد تا متن‌های خواندن مشتمل بر نمونه‌هایی از بندهای موصولی انگلیسی را بخوانند و یادداشت بردارند. در حالی‌که شرکت‌کنندگان در گروه‌های ارتقای متنی داده، درون‌دادهای ارتقایافته را خواندند، شرکت کنندگان در گروه‌های نبود ارتقای متنی داده، درون‌دادهای ارتقانیافته را دریافت کردند. همچنین، شرکت‌کنندگان در گروه‌های تولید برون‌داد زبان‌آموزان ملزم به انجام تکلیف‌های تولید برون‌داد شدند اما از شرکت‌کنندگان در گروه نبود تولید برون‌داد زبان‌آموزان درخواست شد به پرسش‌های درکِ مطلب پاسخ دهند. به دنبال آن، شرکت‌کنندگان در فرآیند تفکر با صدای بلند شرکت کردند. در پایان، آزمون بندهای صفتی انگلیسی دوباره به عنوان پس‌آزمون برگزار شد. یافته‌ها نشان داد که تولید برون‌داد زبان‌آموزان بر توجه شرکت کنندگان و یادگیری بندهای موصولی انگلیسی تأثیری مثبت داشت. هر چند، ارتقای متنی داده فقط بر ارتقای توجه اما نه یادگیری بندهای موصولی انگلیسی تأثیرگذار بود.

کلیدواژه‌ها: توجه، یادگیری، تولید برون‌داد زبان‌آموزان، بند موصولی، ارتقای متنی داده

---

* نویسندهٔ مسئول
[۱] دانشجوی دکتری، گروه زبان و ادبیات انگلیسی، دانشکدهٔ علوم انسانی و اجتماعی، شعبهٔ علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران؛ somsadeghi04@gmail.com
[۲] دانشیار، گروه زبان و ادبیات انگلیسی، دانشکدهٔ علوم انسانی و اجتماعی،شعبهٔ علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران؛ pmaftoon@srbiau.ac.ir
[۳] استادیار، گروه زبان انگلیسی، شعبهٔ گرمسار، دانشگاه آزاد اسلامی، گرمسار، ایران؛ mym1300@gmail.com
شناسه دیجیتال (DOI): 10.22051/lghor.2021.32221.1334

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

کاربرد زبان‌شناسی رایانشی برای پیش‌بینی سطحِ مهارت زبانیِ کتاب‌های درسی زبان‌آموزانِ فارسی

مقاله پژوهشی
صفحات ۵۰-۲۷

مسعود قیومی[۱]

تاریخ دریافت: ۱۳۹۹/۰۵/۲۴ تاریخ پذیرش: ۱۳۹۹/۱۱/۲۱

چکیده

یکی از زیرحوزه‌های ارزیابی مهارت زبانی، پیش‌بینی سطح مهارت زبانی است. این پژوهش به‌دنبال ارائهٔ یک مدل زبان‌شناسی رایانشی برای پیش‌بینی سطح مهارت زبانی و کشف ویژگی‌های کلی این سطوح است. برای این هدف، نگارنده یک پیکرهٔ زبانی از کتاب‌های درسی زبان‌آموزان فارسی گردآوری شده و ویژگی‌های زبان‌شناختی و آماری از این پیکرهٔ متنی برای آموزش سه دسته‌بند به‌عنوان یادگیرنده (زبان‌آموز) استخراج شده‌است. عملکرد مدل‌ها بر مبنای الگوریتم یادگیری و مجموعهٔ ویژگی‌هایی که برای آموزش مدل‌ها به کار رفته‌است متفاوت است. برای ارزیابی مدل‌ها، چهار معیار اندازه‌گیری استاندارد به نام درستی، دقت، فراخوانی و معیار اف استفاده شده‌است. بر اساس یافته‌ها، هنگامی‌که ویژگی‌های آماریِ استخراج شده از متن خام، به کار گرفته شد، مدل ساخته‌شده توسط دسته‌بند جنگل تصادفی بهترین عملکرد مدل را به‌دست آورد. دسته‌بند ماشین بردار پشتیبان با استفاده از ویژگی‌های زبان‌شناختی استخراج‌شده از پیکرهٔ نشانه‌گذاری‌شده به صورت خودکار، بهترین عملکرد مدل را به‌دست آورد. نتایج به‌دست‌آمده مشخص می‌کند که غنی‌سازی مدل و فراهم‌آوری اطلاعات مختلف، تضمین نمی‌کند که یک دسته‌بند (زبان‌آموز) بهترین عملکرد مدل را به‌دست آورد. برای کشف روش‌شناختی آموزشی پنهان در کتاب‌های درسی، عملکرد کلی دسته‌بندها باتوجه به سطح زبانی و دانش زبان‌شناختی به‌کاررفته برای ساخت مدل بررسی شد. براساس یافته‌های به‌دست‌آمده، تعداد ویژگی‌های استخراج‌شده نقش به‌سزایی در آموزش یک دسته‌بند دارد. همچنین، بهترین عملکرد میانگین دسته‌بندها از الگوهای دستوری در سطح مهارت زبانی الف (مبتدی)، به همهٔ دانش زبان‌شناسی در سطوح ب (متوسط) و ج (پیشرفته) گسترش می‌یابد.

کلیدواژه‌ها: یادگیری ماشین، دسته‌بندی، ویژگی، مدل شناختی رایانشی، زبان‌آموز فارسی

[۱] استادیار، پژوهشکده زبان‌شناسی، پژوهشگاه علوم انسانی و مطالعات فرهنگی، تهران، ایران؛ M.Ghayoomi@ihcs.ac.ir
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.32656.1354

مجله افق‌های زبان
LGHOR
دانشگاه الزهرا

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

به کارگیری ارزیابی جایگزین برای تقویت توانایی‌های سازماندهی و کاربردشناختی در تکلیف‌های شفاهی: بررسی بافت انگلیسی به عنوان زبان خارجی

مقاله پژوهشی
صفحات ۷۳-۵۱

فاطمه همتی[*۱]
محبوبه مرتضوی[۲]
روح الله شریفی[۳]

تاریخ دریافت: ۱۳۹۹/۰۹/۰۲ تاریخ پذیرش: ۱۳۹۹/۰۲/۱۴

چکیده

پژوهش حاضر برای بررسی اینکه چگونه دانش سازماندهی و کاربردشناختی می‌تواند در نقش عملکردهای تکنیک‌های ارزیابی در کلاس درس به کار رود، انجام شد. برای این منظور، تأثیرات تکنیک‌های معلم، همکلاسی و خود-ارزیابی که بر فراهم کردن بازخورد دربارهٔ دانش سازماندهی و کاربردشناختی دانش‌آموزان تمرکز داشتند، مقایسه شدند. تعداد کل ۹۸ زبان‌آموز دختر ایرانی که در سطح پیش- متوسط در یک موسسهٔ زبان در تهران تحصیل می‌کردند، در این پژوهش شرکت کردند. آن‌ها به‌طور تصادفی به یکی از سه گروه خودارزیابی (تعداد=۳۲)، همکلاسی-ارزیابی (تعداد=۳۳) و معلم-ارزیابی (تعداد=۳۳) دسته‌بندی شدند. توانایی کاربردشناسی با استفاده از نقش بازی کردن‌های شفاهیِ گفتمان و دانش سازماندهی با مقیاسی که دانش دستوری و متنی شرکت‌کنندگان را می‌سنجید، ارزیابی شدند. یافته‌های آزمون واریانس یک‌طرفه نشان داد که خود-ارزیابی سپس ارزیابی با همکلاسی، بر تکنیک ارزیابی معلم در ارتقای توانایی سازماندهی و کاربردشناختی برتری داشت. به طور کلی، یافته‌های به‌دست‌آمده تأییدی برای به‌کارگیری روش‌های جایگزینِ ارزیابی به مثابه ابزارهای آموزشی برای تقویت توانایی زبانی به شمار می‌آید. کاربردهای پژوهش برای آموزش زبان انگلیسی به عنوان زبان خارجی موردِ بحث قرار گرفتند.

کلیدواژه‌ها: خود-ارزیابی، همکلاسی-ارزیابی، معلم-ارزیابی، دانش کاربردشناختی، دانش سازماندهی

---

* نویسندهٔ مسئول
[۱] دانشیار، گروه آموزش زبان انگلیسی، دانشگاه پیام نور، تهران، ایران؛ f.hemmati@jonoub.tpnu.ac.ir
[۲] دانش‌آموخته دکتری، گروه آموزش زبان انگلیسی، دانشگاه پیام نور، تهران، ایران؛ Mortazavi.0m@gmail.com
[۳] استادیار، گروه آموزش زبان انگلیسی، دانشگاه پیام نور، تهران، ایران؛rsharifi@pnu.ac.ir
شناسه دیجیتال (DOI): 10.22051/lghor.2021.32932.1361

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

بررسی هوش فرهنگی و هوش عاطفی به عنوان پیش‌بینی‌کننده‌های فرسودگی شغلی مابین معلمان زبان انگلیسی عراقی

مقاله پژوهشی
صفحات ۹۶-۷۵

رضا باقری نویسی[*۱]
زینب الأسدی[۲]

تاریخ دریافت: ۱۳۹۹/۰۷/۲۳ تاریخ پذیرش: ۱۳۹۹/۱۱/۲۱

چکیده

پژوهش حاضر، تلاشی برای کشف ارتباط بین فرسودگی شغلی معلمان زبان انگلیسی به عنوان زبان خارجی و هوش فرهنگی و عاطفی آن‌ها است. افزون بر این، هدف، کشف این نکته بود که کدام متغیر، هوش هیجانی یا فرهنگی، می‌تواند پیش‌بینِ بهتری در مورد فرسودگی شغلی معلم باشد. این مطالعه همچنین به جستجوی رابطهٔ محتمل جنسیت و فرسودگی معلمان زبان انگلیسی عراقی پرداخته و بررسی کرد که کدام جنسیت بیشتر تحت تأثیر فرسودگی شغلی قرار گرفته‌است. در نهایت، معلمان مرد و زن عراقی با توجه به هوش هیجانی و هوش فرهنگی مورد سنجش و مقایسه قرار گرفتند. به این منظور، سه پرسشنامه ۱) پرسشنامهٔ هدفِ فرسودگی شغلی ماسلاچ، ۲) مقیاس هوش فرهنگی که توسط ون داین و همکاران (۲۰۰۸) طراحی شده، و ۳) مقیاس هوش عاطفی وانگ و لا، برای بیش از ۲۰۰ معلم، یا حضوری یا از طریق ایمیل و یا از طریق لینک از پیش ساخته‌شده گوگل‌داک از سه پرسشنامه ارائه شد. برخی معلمان از به اشتراک گذاشتن پاسخ‌های خود بی‌میل بودند و از پر کردن پرسشنامه پرهیز کردند، ولی ۱۶۴ معلم زبان انگلیسی عراقی هر سه پرسشنامه را پر کردند. یافته‌ها نشان داد که معلمان زبان انگلیسی عراقی و هوش فرهنگی و فرسودگی شغلی آن‌ها، به صورت معنادار اما منفی با هم همبستگی داشت. هوش هیجانی و فرهنگی مشابه هم و به طور معناداری پیش بینی کننده فرسودگی شغلی بودند. همچنین، فرسودگی شغلی و جنسیت به صورت معناداری همبستگی داشتند و معلمان زبان انگلیسی عراقی زن بیشتر مستعد و در معرض فرسودگی شغلی بودند. یافته‌ها به طور ضمنی اشاره دارد که معلمان آگاه از نظر فرهنگی و عاطفی ممکن است بتوانند با اقدامات پیشگیرانه، فرسودگی شغلی خود را کنترل کرده و یا کاهش دهند.

کلیدواژه‌ها: فرسودگی شغلی، هوش عاطفی، هوش فرهنگی، معلمان زبان انگلیسی به عنوان زبان خارجی، جنسیت

---

* نویسندهٔ مسئول
[۱] دانشیار، گروه زبان و ادبیات انگلیسی، دانشگاه قم، قم، ایران؛ r.bagherinevisi@qom.ac.ir، re.baghery@gmail.com
[۲] دانش‌آموختهٔ کارشناسی ارشد، گروه زبان و ادبیات انگلیسی، دانشگاه قم، قم، ایران؛ zalasadi92@yahoo.com
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.33516.1383

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

بررسی انواع مذاکره فراگیران زبان انگلیسی ایرانی در نگارش با الگوهای یادگیری حمایتی متفاوت

مقاله پژوهشی
صفحات ۱۱۴-۹۷

عایشه محمدزاده[۱]
توران آهور*[۲]
مهناز سعیدی[۳]

تاریخ دریافت: ۱۳۹۸/۱۱/۲۸ تاریخ پذیرش: ۱۳۹۹/۰۷/۰۵

چکیده

رویکردهای زبان آموز-محور در فراگیری زبان دوم و رویکرد فرایند در آموزش نگارش زبان دوم، توجه شماری از پژوهشگران را نسبت به اظهارنظرهای زبان‌آموزان در پشت سرنهادن پیش‌نویس‌ها و بازبینی‌های چندگانه برانگیخته است. این پژوهش، بر اساس مفهوم یادگیری حمایتی در نظریهٔ اجتماعی فرهنگی ویگوتسکی و رویکرد فرایند به نگارش زبان دوم، انواع مذاکرهٔ فراگیران زبان انگلیسی ایرانی در گروه‌های کوچک را با الگوهای یادگیری حمایتی؛ متقارن و نا متقارن بررسی کرد. به این منظور، ۱۵ زبان‌آموز در سطح‌های مهارت زبانی سطح متوسط بالا و پائین به سه گروه با الگوهای حمایتی متفاوت دسته‌بندی شدند: یک گروه نامتقارن با دو زبان‌آموز سطح متوسط بالا -سه زبان‌آموز سطح متوسط پائین، گروه‌های متقارن دوم با پنج زبان‌آموز سطح متوسط بالا و گروه دیگر با پنج زبان‌آموز متقارن سطح متوسط پائین. رونوشت‌ها برای تشخیص انواع مذاکره به وسیلهٔ عملکردهای زبانی در بین گروه‌های مختلف، تجزیه و تحلیل شدند. بنابراین، دو مقوله‌های اصلی، *پاسخ* و *درخواست* و زیرمقوله‌های آن‌ها یافت شد. زیرمقوله‌های «موافقت»، «توضیح»، «نظر دادن»، «آموزش»، «از نو گفتن» و «پیشنهاد دادن» مربوط به نخستین مجموعه اصلی و «بررسی درک مطلب»، «استنباط نظرات» و «سوال پرسیدن» مربوطه به مقوله اصلی دوم بودند. بر مبنای یافته‌های آزمون خی‌دو، انواع مذاکره ارتباط معناداری با الگوی یادگیری حمایتی دارد به طوری‌که گروه نامقارن (بالا و پائین) در تعداد عملکردهای زبانی مورد استفاده بر گروه‌های متقارن (پائین، پائین؛ بالا، بالا) برتری داشت. با این وجود، افراد در همهٔ گروه‌ها از کیفیت بالا و تقابل در تعامل بهره‌مند شدند. بر اساس یافته‌ها به معلم‌ها پیشنهاد می‌شود آگاهی زبان‌آموزان دربارهٔ نقاط قوت و توانایی‌های گوناگون که الگوهای متفاوت یادگیری حمایتی به آن‌ها می‌دهد را بالا ببرند.

کلیدواژه‌ها: حمایت نامتقارن، نگارش مشارکتی، فراگیران زبان انگلیسی به عنوان زبان خارجی، انواع مذاکره، حمایت متقارن

---

* نویسندهٔ مسئول
[۱] دانشجوی دکتری، گروه زبان انگلیسی، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران؛ Mohammadzadeh79@gonbadiau.ac.ir
[۲] استادیار، گروه زبان انگلیسی، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران؛ touranahour@yahoo.com
[۳] دانشیار، گروه زبان انگلیسی، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران؛ m_saeidi@iaut.ac.ir
شناسه دیجیتال (DOI): 10.22051/LGHOR.2020.30357.1259

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

درک آموزشی معلمان زبان انگلیسی ایرانی از نقش ادغام فناوری در کلاس‌های درسی

مقاله پژوهشی
صفحات ۱۳۷-۱۱۵

سمیرا عرب زمانی[۱]
علی عرب مفرد[*۲]
علی رجبی استرآبادی[۳]

تاریخ دریافت: ۱۳۹۹/۰۸/۲۰ تاریخ پذیرش: ۱۴۰۰/۰۲/۰۴

چکیده

اگرچه پژوهش‌های گسترده‌ای روی ادغام موفقیت‌آمیز فناوری، به اجرا درآمده‌است، اما به نظر می‌رسد هنوز در کلاس، گنجاندنِ فناوریِ قابلِ دسترس، مهیاکردن محیطی فناوری‌محور و به کار بردن فناوریِ سطح بالا زیاد مورد استفاده قرار نمی‌گیرد. همچنین، ادغام فناوری موانع دیگری شامل دانش فناوری‌محور دانش‌آموزان و معلمان و مشکلات مالی مدارس را نشان می‌دهد. برخلاف یافته‌های پیشین در بافت‌های زبان انگلیسی به عنوان زبان خارجی، به نظر می‌رسد شواهد اندکی در مورد کاربردهای فناوری کلاس معلمان ایرانی در دسترس باشد. بنابراین، این مقاله بر آن است تا درکِ آموزشی معلمان ایرانی دربارهٔ ادغام فناوری و روشن ساختن چگونگی تأثیرات ادغام فناوری بر عملکرد کلاس‌های معلمان ارائه دهد. در این راستا، با ۱۰ معلم زبان انگلیسی ایرانی در مدارس خصوصی و دولتی در استان گلستان مصاحبه شد. داده‌های کیفی مصاحبه‌ها به انگلیسی ثبت و رونویسی شد؛ آن‌ها سپس، با استفاده از تجزیه و تحلیل موضوعی (براون و کلارک، ۲۰۰۶) از طریق نرم‌افزار رایانه‌ای ان ویوو ۱۲، کدگذاری و بررسی شد تا ابعاد و مؤلفه‌های ادغام مدل مفهومی فناوری را نشان دهد. یکی از موضوعاتی که از تحلیل موضوعی درکِ معلمان ایرانی استخراج شد، نبود دانش فنی بود. موضوع دیگری که به این کاربردِ محدودِ فناوری اشاره می‌کند، استفاده کردن معلمان از رایانه به عنوان ابزاری تکمیلی است. افزون بر این، این یافته به دست آمد که بافت آموزشی شرایط فناوری نیاز-محور را برای آن‌ها ایجاد نمی‌کند که منجر به استفادهٔ غیر مؤثر معلمان از فناوری می‌شود.

کلیدواژه‌ها: درک معلم، ادغام فناوری، تحلیل موضوعی، نظریهٔ زمینه‌ای، معلمان زبان انگلیسی ایرانی

---

* نویسندهٔ مسئول
[۱] دانشجوی کارشناسی ارشد، گروه آموزش زبان انگلیسی، دانشگاه علوم انسانی، واحد گرگان، دانشگاه آزاد اسلامی، گرگان، ایران؛ teflsaz@gmail.com
[۲] استادیار، گروه آموزش زبان انگلیسی، دانشگاه علوم انسانی، واحد گرگان، دانشگاه آزاد اسلامی، گرگان، ایران؛ a.arabmofrad@gu.ac.ir
[۳] دانشجوی کارشناسی، گروه آموزش زبان انگلیسی، دانشگاه علوم انسانی، واحد گرگان، دانشگاه آزاد اسلامی، گرگان، ایران؛ alirajabi167@yahoo.com
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.33557.1398

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

ریشه‌یابی خطاهای مربوط به مطابقتِ شمار در فراگیران سطح متوسط-پیشرفتهٔ زبان انگلیسی

مقاله پژوهشی
صفحات ۱۶۶-۱۳۹

ابراهیم صفایی[۱]

تاریخ دریافت: ۱۳۹۹/۱۱/۱۳ تاریخ پذیرش: ۱۳۹۹/۱۲/۱۶

چکیده

این پژوهش، میزانِ حساسیتی که زبان‌آموزان بزرگسال انگلیسی به خطاهای مطابقت فاعل-فعل در فعل‌های موضوعی و ربطی دارند، را بررسی می‌کند. یک گروه از فراگیران زبان دوم متوسط- پیشرفته ایرانی و گروهی از گویشوران بومی انگلیسی در فعالیت مقبولیت قضاوتِ سریع برخط شرکت کردند. یافته‌ها از این قرار است: درحالی‌که گویشوران بومی انگلیسی به خطاهای مطابقت شمار در هر دو نوع فعل حساس بودند، فراگیران زبان دوم به آن‌ها حساس نبودند. برای گروه فراگیران زبان دوم، یافته‌ها خطاهای نظام‌مند با الگوی زیر را نشان می‌دهد. برای خطاهای مطابقت شمار در فعل‌های موضوعی، عدم تقارنِ حذف-اضافه وجود داشت؛ در حالی‌که فراگیران زبان دوم به خطاهای حذف حساس نبودند، آن‌ها به خطاهای اضافه کردن در صورت‌های ایستا حساس بودند. برای خطاهای مطابقت شمار در فعل های ربطی‌ها، عدم تقارن مفرد-جمع وجود داشت که بیشترین مقدار خطا در آرایش‌های فاعل جمع + فعل ربطی مفرد در مقایسه با آرایش‌های فاعل مفرد + فعل ربطی جمع بود. با این وجود، به نظر می‌رسد مهارت زبانی یک پیش‌بین‌کنندهٔ قوی در حساسیت فراگیران زبان دوم به خطاهای مطابقت باشد. یافته‌ها، فرضیه ابهام واژگانی (مک کارتی، ۲۰۰۷؛ ۲۰۰۸؛ ۲۰۱۲) را تأیید می‌کند.

کلیدواژه‌ها: خطاهای مغایرت مشخصه، عدم تقارن حذف-اضافه، عدم تقارن مفرد-جمع، مطابقت فاعل و فعل، خطاهای ابهام واژگانی

[۱] استادیار، گروه زبان انگلیسی، دانشگاه آزاد اسلامی، واحد لامرد، لامرد، ایران؛
es.safaie@iaulamerd.ac.ir ،safaie124@yahoo.com
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.34969.1438

مجله افق‌های زبان
LGHOR
دانشگاه الزهرا

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

تأثیر یادگیری وارونه بر مهارت نگـارش انگلیـسی و خودکارآمدی دانشجویان پزشکی ایرانی

مقاله پژوهشی
صفحات ۱۸۶-۱۶۷

سپیده نوری‌نژاد[۱]
احسان هادی‌پور فرد[*۲]
محمد باولی[۳]

تاریخ دریافت: ۱۳۹۹/۰۶/۱۶ تاریخ پذیرش: ۱۳۹۹/۱۰/۱۳

چکیده

پژوهش حاضر، تأثیر آموزش وارونه بر خودکارآمدی نگارش و عملکرد نگارش دانشجویان پزشکی در دانشگاه علوم پزشکی شیراز را بررسی می‌کند. پنجاه دانشجو در این پژوهش تجربی شرکت کردند. آن‌ها به عنوان گروه آزمودنی مشتمل بر ۲۵ دانشجو و گروه کنترلِ ۲۵ دانشجو دسته‌بندی شدند. گروه کنترل با روش آمـوزش نگـارش سـنتی (غیرِ وارونه) آموزش دیدند، در حالی‌که گروه آزمایش با شیوهٔ آموزش وارونه آموزش داده شدند. پرسشنامه به کـار رفته در این مطالعه، از پژوهش پریکلز (۱۹۹۴) اقتباس شد و بر اساس مقیاس لیکرت برای نگارش خودباوری نمره داده شد. داده‌ها در یک طرح پیش‌آزمون-پس‌آزمون گردآوری شدند. یافته‌ها نشان داد کـه آمـوزش وارونـه تـأثیر مثبت بیشتری بر بهبود هر دویِ خودکارآمدی نگارش و عملکرد نگارشی زبان‌آموزان در مقایسه با آن‌هایی کـه بـه روش سنتی آموزش دیده بودند، داشت. یافته‌ها در این بررسی، دیدگاه معلمان آموزش زبـان انگلیـسی بـه عنـوان زبان خارجی برای فهم روش‌های نوین آموزشی را گسترش داد.

کلیدواژه‌ها: آموزش سنتی، آموزش وارونه، عملکرد نگارش، خودکارآمدی، خودکارآمدیِ نگارش

---

* نویسندهٔ مسئول
[۱] دانشجوی دکتری، گروه زبان انگلیسی، شعبهٔ شیراز، دانشگاه آزاد اسلامی، شیراز، ایران؛
[۲] استادیار، گروه زبان انگلیسی، شعبهٔ شیراز، دانشگاه آزاد اسلامی، شیراز، ایران؛ e.ehadipour@gmail.com
[۳] گروه زبان انگلیسی، شعبهٔ شیراز، دانشگاه آزاد اسلامی، شیراز، ایران؛ mbvl57@gmail.com
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.34132.1409

مجله افق‌های زبان
LGHOR
دانشگاه الزهرا

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

تأثیر پیاده‌سازی مدل کلاس وارونه بر تفکر انتقادی و عملکرد فراگیران زبان انگلیسی ایرانی در یادگیری دستور زبان

مقاله پژوهشی
صفحات ۲۰۳-۱۸۷

زهرا افضلی[۱]
سیروس ایزدپناه[*۲]

تاریخ دریافت: ۱۳۹۹/۰۶/۱۶ تاریخ پذیرش: ۱۳۹۹/۱۰/۱۳

چکیده

کلاس وارونه، یک رویکرد آموزشی است که می‌تواند سببِ ارتباطات رودرروی اضافی با دانش‌آموزان در کلاس شود. اگر چه پژوهش‌هایی وجود دارد که اغلب کلاس وارونه را به طور کلی بررسی کرده‌اند، اما مطالعات اندکی به ویژه تفکر انتقادی و عملکرد در یادگیری دستور زبان را بررسی کرده‌اند. طرح این پژوهش نیمه‌تجربی بود. این مطالعه در یک کلاس انگلیسی به مدت شش هفته در سال ۲۰۲۰ به اجرا درآمده‌است. ۳۶۰ زبان‌آموز انگلیسی با استفاده از نمونه‌گیری خوشه‌ای چندمرحله‌ای از دو موسسهٔ مختلف در محدوده سنی ۱۳تا سالگی ۱۹ که در سطوح متوسط و بالاتر از آن تحصیل می‌کردند، انتخاب شدند. از میان پنجاه فیلم پیرامونِ دستور، هشت محتوای دستور زبان بر اساس شاخص روایی محتوا و نسبت روایی محتوا انتخاب شدند. پرسشنامه‌های آزمون‌های عملکرد انگلیسی و تفکر انتقادی به کار گرفته شدند. گروه آزمایشی با مدل کلاس درسی وارونه آموزش دیدند در حالی که دوره‌های آموزشی بر اساس برنامه فعلی در گروه کنترل به اجرا درآمد. معلوم شد که هیچ اختلاف چشمگیری میان نمره‌های پیش‌آزمون و پس‌آزمون گروه کنترل وجود ندارد، در حالی که بین نمره‌های پیش‌آزمون و پس‌آزمون گروه آزمایش تفاوت معناداری وجود دارد. می‌توان به معلمان پیشنهاد داد از مدل کلاس وارونه استفاده کنند تا تفکر انتقادی و عملکرد فراگیران را تقویت کنند.

کلیدواژه‌ها: تفکر انتقادی، کلاس وارونه، یادگیری وارونه، آموزش وارونه، عملکرد

---

* نویسندهٔ مسئول
[۱] دانشجوی کارشناسی ارشد، گروه آموزش زبان انگلیسی، واحد زنجان، دانشگاه آزاد اسلامی، زنجان، ایران؛ afzalizahra90@gmail.com
[۲] استادیار، گروه آموزش زبان انگلیسی، واحد زنجان، دانشگاه آزاد اسلامی، زنجان، ایران؛ sirosizadpanah@iauz.ac.ir
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.32997.1362

مجله افق‌های زبان
LGHOR
دانشگاه الزهرا

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

تأثیر ارزیابی پویای گروهی رایانشی بر درک مطلب شنیداری زبان آموزان انگلیسی به عنوان زبان خارجی در دو جنسیت

مقاله پژوهشی
صفحات ۲۲۵-۲۰۵

شاهین عباسی دلوند[۱]
داوود مشهدی حیدر[۲*]

تاریخ دریافت: ۱۳۹۹/۱۱/۱۵ تاریخ پذیرش: ۱۴۰۰/۰۱/۳۱

چکیده

در چارچوب نظریه اجتماعی-فرهنگی ویگوتسکی، این پژوهش قصد دارد به بررسی اینکه آیا ارزیابی پویای گروهی رایانشی، توانایی درک مطلب شنیداری زبان‌آموزان مرد و زن ایرانی را تحت تأثیر قرار داده‌است، بپردازد. داده‌ها از طریق برگزاری پیش‌آزمون و پس‌آزمون‌های درک مطلب شنیداری مابین ۱۴۰ شرکت‌کننده که به فراگیران مرد و زن در گروه‌های کنترلی و آزمایشی دسته‌بندی شده بودند، گردآوری شد. در هرگروه از مطالعه، ۳۵ زبان‌آموز زن و ۳۵ زبان‌آموز مرد بودند. شرکت‌کنندگان در گروه‌های آزمایش، در معرض ارزیابی پویایی گروهی قرار گرفتند تا به صورت تعاملی، تکلیف‌های منتخب درک مطلب شنیداری را انجام دهند و معلم نیز حمایت‌های لازم را هم فراهم نمود. تجزیه و تحلیل کمی پیش‌آزمون‌ها و پس‌آزمون‌های درک مطلب شنیداری میان گروه‌های زن و مرد از طریق فرآیند تحلیل واریانس وکوواریانس دوسویه انجام گرفت. یافته‌ها نشان دادند که فراگیران مرد و زن هر دو، در گروه‌های آزمایشی به طور معناداری عملکرد بهتری از فراگیران در گروه‌های کنترل داشتند. اما تفاوت معناداری میان توانایی درک مطلب شنیداری گروه‌های جنسیت در گروه‌های آزمایشی مشاهده نشد. یافته‌ها به بهره‌گیری مؤثرِ ارزیابی پویای گروهی از طریق نرم‌افزار به منظور بهبود توانایی درک مطلب شنیداری زبان‌آموزان کمک کردند، به این معنا که پیشنهاد می‌شود معلم‌ها از ابزارهای فناوری آگاه باشند تا بتوانند برای زبان‌آموزان بستری تعاملی جهتِ پیشرفت مهارت‌ها و زیر مهارت‌های زبانی‌شان فراهم نمایند.

کلیدواژه‌ها: آموزش زبان به کمک رایانه، ارزیابی پویای گروهی، انگلیسی به عنوان زبان خارجی، توانایی درک شنیداری، منطقهٔ مجاور رشد

---

* نویسندهٔ مسئول
[۱] دانشجوی دکتری، گروه زبان انگلیسی، واحد تنکابن، دانشگاه آزاد اسلامی، تنکابن، ایران؛ Shahin.abacy@gmail.com
[۲] استادیار، گروه زبان انگلیسی، واحد تنکابن، دانشگاه آزاد اسلامی، تنکابن، ایران؛ d.mashhadiheydar@toniau.ac.ir
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.34839.1440

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

تأثیر تعامل معلم- دانشجو و خود-پنداری تحصیلی بر پیشرفت تحصیلی فراگیران زبان انگلیسی

مقاله پژوهشی
صفحات ۲۳۷-۲۲۷

نفیسه اسدزاده ملکی[۱]
مسعود ذوقی[۲*]
نادر اسدی آیدین‌لو[۳]

تاریخ دریافت: ۱۳۹۹/۱۰/۰۷ تاریخ پذیرش: ۱۴۰۰/۰۲/۲۲

چکیده

در حالی که پژوهش‌های پیشرفت تحصیلی پیشین، طیف گسترده‌ای از متغیرهای شناختی را بررسی کرده‌اند، متغیرهای غیر شناختی خاصی مانند تعامل معلم-دانشجو و خودپنداره تحصیلی در رابطه با پیشرفت تحصیلی هنوز نیازمند بررسی هستند. به این منظور، از یک روش تحقیق ضریب همبستگی برای مطالعهٔ خودپنداره تحصیلی و ویژگی‌های تعامل بین معلم و دانشجو استفاده شده‌است که به عنوان عواملی که نقش مهمی در پیشرفت تحصیلی زبان آموزان ایفا می‌کنند، شناخته شده‌اند. شرکت‌کنندگان در این مطالعه شامل ۲۱۸ دانشجوی زبان انگلیسی ایرانی با دامنهٔ سنی ۱۸ تا ۴۵ سال بودند. رویه‌های آزمون-فرضیه مدل معادلات ساختاری نشان داد که بین مدل و داده‌های به‌دست‌آمده تناسب منطقی وجود دارد ($\chi^2 = 0.37$, GFI = 0.99, CFI = 0.99, RMSEA = 0.01). اثبات شد که میزان روابط بین معلم و دانشجو به صورتی معنادار و مستقیم بر پیشرفت تحصیلی دانشجویان انگلیسی به عنوان زبان خارجی تأثیرگذار بود ($r = 0.35, p = 0.45$). افزون بر این، خودپنداره تحصیلی، عامل مهمی در بهبود پیشرفت دانشجویان شناخته شد ($r = 0.24, p = 0.32$). یافته‌های به‌دست‌آمده همچنین نقش میانجیگری خودپنداره تحصیلی زبان‌آموزان در ارتقای پیشرفت تحصیلی زبان‌آموزان از طریق افزایش تعامل بین دانشجو و معلم آن‌ها را نشان داد. نقش بنیادین ویژگی‌های جسمی-روانی، شامل روابط معلم و دانشجو و خودپنداره تحصیلی در پیشرفت آموزشی دانشجویان زبان خارجی تأیید شد. یافته‌های این پژوهش تجربی، کاربردهایی برای موقعیت‌های آموزشی دارد.

کلیدواژه‌ها: پیشرفت تحصیلی، خود پنداره تحصیلی، روش همبستگی، مدل معادلات ساختاری، تعامل معلم-دانشجو

---

* نویسندهٔ مسئول
[۱] دانشجوی دکترا، گروه زبان انگلیسی، دانشکده علوم انسانی، واحد اهر، دانشگاه آزاد اسلامی، اهر، ایران؛
Iran.nafiseh.asadzadeh@malekaniau.ac.ir
[۲] استادیار، گروه زبان انگلیسی، دانشکده علوم انسانی، واحد اهر، دانشگاه آزاد اسلامی، اهر، ایران (نویسندهٔ مسئول)؛
m-zoghi@iau-ahar.ac.ir،dr.m.zoghi@gmail.com
[۳] استادیار، گروه زبان انگلیسی، دانشکده علوم انسانی، واحد اهر، دانشگاه آزاد اسلامی، اهر، ایران؛ naderasadi@yahoo.com
شناسه دیجیتال (DOI): 10.22051/LGHOR.2021.34581.1426

مجله افق‌های زبان
LGHOR
دانشگاه الزهرا

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

نقش تداعی معنایی در ابهام‌زدایی جمله‌های موصولی در زبان فارسی

مقاله پژوهشی
صفحات ۲۵۲-۲۳۹

اسکندر صمدی[۱]
محمدنبی کریمی[۲*]
عصمت بابایی[۳]

تاریخ دریافت: ۱۳۹۹/۰۳/۳۰ تاریخ پذیرش: ۱۳۹۹/۰۹/۱۵

چکیده

پژوهشِ حاضر، به بررسی نقش تداعیِ معنایی در پردازش جمله‌های مبهمِ دربردارندهٔ بندهای موصولی که گویشوران فارسی قبل از آن یک گروه اسمی مرکب مبهم به کار می‌برند، پرداخته است. به این منظور، در یک تکلیف خواندن خودآموز با استفاده از نرم افزار ای. پرایم، ۶۳ گویشور فارسی، جمله‌های دربردارندهٔ بندهای موصولی مبهم در زبان اولشان یعنی فارسی را خواندند. نوع رابطه معنایی در این پژوهش، از نوع رابطهٔ بین بند موصولی و یکی از گروه‌های اسمی مرکب بود تا کشف شود آیا تداعیِ گروه‌های اسمی از طریق این آزمایش معنایی، اولویت‌های الحاقی گویشوران فارسی را متأثر می‌کند. یافته‌های تفسیری پست‌های آفلاین (اولویت‌های الحاقی بندهای موصولی) و داده‌های برخط (مدت‌زمان‌های خواندن) نشان داد که تداعی معنایی، اولویت‌های الحاقی شرکت‌کنندگان را متأثر می‌سازد که بدین معناست که اولویت‌های تجزیه آن‌ها فقط به وسیلهٔ اطلاعات نحوی هدایت نمی‌شوند. این یافته‌ها همسو با مدل‌های محدودیت-بنیاد تجزیه جمله‌ها است که در این مدل ها تصور بر آن است که در هنگام تجزیه جمله‌ها، چندین منبع اطلاعاتی با یک‌دیگر تعامل دارند و هر یک از آن‌ها تفسیر را به روش خاصی محدود می‌سازد. همچنین، یافته‌ها پیش‌بینی‌های مدل فعال‌سازی گسترشی و نظریهٔ تجزیه واژگانی را پشتیبانی می‌کند.

کلیدواژه‌ها: اولویت الحاقی بند موصولی، تفکیک‌پذیری ابهام، معناشناسی، تداعی معنایی، گویشوران فارسی

---

* نویسندهٔ مسئول
[۱] دانشجوی دکتری زبان‌شناسی کاربردی، گروه زبان‌های خارجی، دانشگاه خوارزمی، تهران، ایران؛ Std_samadi@khu.ac.ir
[۲] دانشیار زبان‌شناسی کاربردی، گروه زبان‌های خارجی، دانشگاه خوارزمی، تهران، ایران؛ Karimi_mn@khu.ac.ir
[۳] استاد زبان‌شناسی کاربردی، گروه زبان‌های خارجی، دانشگاه خوارزمی، تهران، ایران؛ Babaii@khu.ac.ir
شناسه دیجیتال (DOI): 10.22051/LGHOR.2020.31707.1323

دوفصلنامهٔ علمی افق‌های زبان دانشگاه الزهرا(س)
سال ششم، شمارهٔ ۱، بهار و تابستان ۱۴۰۱

فهرست مطالب

معاونت پژوهشی دانشگاه

دوفصلنامهٔ علمی

مجلهٔ افق‌های زبان دانشگاه الزهرا(س)

سال ششم، شماره اول، بهار و تابستان ۱۴۰۱ (دوفصلنامه ـ پیاپی ۱۱)

صاحب امتیاز: **دانشگاه الزهرا(س)**

مدیر مسئول: **دکتر محسن شیرازی‌زاده**

سردبیر: **دکتر الهه ستوده‌نما**

ویراستار زبان انگلیسی: **دکتر شهره تیمورنژاد**

ویراستاران زبان فارسی: **دکتر نرجس منفرد**

مدیر اجرایی: **نرگس جعفری**

اعضای هیئت تحریریه

آدرس: تهران، ونک، دانشگاه الزهرا(س)، دانشکدهٔ ادبیات، کد پستی: ۱۹۹۳۸۹۱۱۷۶

پست الکترونیکی: lghorizons@alzahra.acir

شاپا: ۳۵۰x-۲۵۸۸

شاپای الکترونیکی: ۵۶۳۴-۲۵۸۸

مجلهٔ افق های زبان با همکاری انجمن زبان شناسی ایران منتشر می‌شود.

انجمن زبان‌شناسی ایران

LGHOR