Interdisciplinary Studies in English Language Teaching

ISELT

Editorial Board

UNIVERSITY OF
MAZANDARAN

About the Journal

Interdisciplinary Studies in English Language Teaching (ISELT), published by the English Department at the University of Mazandaran, Babolsar, Iran, provides a platform for cutting-edge research in English Language Teaching (ELT). The journal focuses on diverse themes such as ELT methodologies, language acquisition, curriculum design, and technology-enhanced learning, as well as applied linguistics, language testing, and teacher professional development, encouraging submissions that adopt various research methods, including empirical studies, theoretical papers, and action research. As a double-blind peer-reviewed and open-access journal, ISELT aims to advance scholarship in areas like bilingualism, multilingual education, and critical perspectives in ELT, welcoming contributions with both local and global relevance. The journal is published semiannually in electronic format, allowing readers to access, distribute, and link to the full texts without needing prior permission. All submissions are checked for similarity using iThenticate to ensure academic integrity.

Guide for Authors

Manuscript Submission Guidelines

Thank you for choosing to submit your paper to the *Interdisciplinary Studies in English Language Teaching (ISELT)*. Please take the time to read and follow the manuscript submission guidelines, ensuring that your paper is formatted according to ISELT's requirements. Authors are required to provide a Username/Password for the ISELT journal. Manuscripts should be submitted through our online submission system; please do not send your manuscripts by e-mail unless specifically requested. The receipt of the manuscript will be acknowledged via e-mail. Failure to follow the manuscript preparation and submission guidelines may result in delays in the manuscript processing.

Manuscripts must be written in English (either American or British), with consistent spelling throughout the text. Translated articles will not be accepted. Authors are strongly advised to check their manuscripts for plagiarism and fraudulent data. The ISELT editorial team uses both reviewer checks and a plagiarism prevention tool (iThenticate) for all submissions. At least two reviewers are invited to review each submitted manuscript.

Authors should kindly note that:

ISELT manuscripts are reviewed only if they:
• are original,
• are not under consideration by any other journal or publisher,
• have not been previously published in whole or in part,
• have not been previously accepted for publication, and
• will not be submitted elsewhere until a decision is made regarding their publication.

The submitted article should include the following files:

The first file: Title page

The title page, as a separate file, should contain the full title of the paper, with the name(s) of all authors, academic level (e.g., Assistant Professor, Associate Professor, Professor, or PhD Candidate), department, university, country, academic email address, and, if available, telephone number. The first and corresponding authors must be clearly indicated. The corresponding author's name should be marked with an asterisk (*). If the article is based on a thesis or dissertation, this should be mentioned on the first page (e.g., "This article is based on the Ph.D. thesis of …"). The title page should be submitted as an editable MS Word file (not a PDF).

نویسندگان فارسی‌زبان لطفاً عنوان مقاله، وابستگی دانشگاهی، و جایگاه علمی نویسندگان را به فارسی هم بنویسند.

Font style and size of the title page:

Document size: A4
Article Title: Times New Roman 14 pt., bold
Author name(s): Times New Roman 12 pt., bold
Affiliation and academic email address: Times New Roman 12 pt.

The second file: Abstract, without authors' names

Document size: A4
Article Title: Times New Roman 14 pt., bold
Abstract: Times New Roman 12 pt., double spacing (between 200 and 250 words)
Keywords: Times New Roman 12 pt. (four or five words)

نویسندگان فارسی زبان لطفاً ترجمه فارسی عنوان مقاله و چکیده را در یک فایل جداگانه بارگذاری نمایند.

The third file: The main article without authors' names

A typical journal article includes the title, abstract, introduction, literature review, methods, results, discussion, (conclusions), and references. The length should be between 6500 and 7500 words, excluding the abstract and references.

Font style and size of the main article without authors' names:

The full title of the article: Capitalize the first letter of the main words, using Times New Roman 12 pt., bold. One line space must be left after the title.

Abstract: Times New Roman 12 pt., single spacing, between 200 and 250 words, including key terms (up to five terms).

Headings (Level 1): **INTRODUCTION**, **LITERATURE REVIEW**, **METHODS**, **RESULTS**, **DISCUSSION** (**CONCLUSIONS**), and **REFERENCES**" must be left-aligned, bold, uppercase, and printed in all capital letters (Times New Roman 11 pt.).

Subheadings (Level 2): Left-aligned, bold, upper/lowercase (only the first letter of the first word capitalized), Times New Roman 11 pt.

Subheadings (Level 3): Left-aligned, bold, upper/lowercase (only the first letter of the first word capitalized), Times New Roman 11 pt., and the same format applies to Levels 4 and 5.

Citations and References:

Manuscripts submitted to ISELT should follow the *American Psychological Association* (APA 6th edition) style, available at: http://www.apa.org/books/4200061.html. References should be listed alphabetically at the end of the paper. It is highly recommended to hyperlink each in-text citation to its corresponding reference entry to facilitate direct access to the cited work for readers.

Copyright Form:

Papers accepted for publication become the copyright of the Editorial Office of the *Interdisciplinary Studies in English Language Teaching (ISELT)*, University of Mazandaran. Authors will be asked to sign a *Publication Ethics form/Copyright Assignment Form (Form 1)*. By signing the copyright transfer, authors are assumed to have obtained permission to use any copyrighted or previously published material. All authors must read and agree to the conditions outlined in the *Copyright Assignment Form*, and must sign the form, or agree that the corresponding author can sign on their behalf. Articles cannot be published until a signed *Copyright Assignment Form* has been received.

Conflict of Interest Form (Form 2):

We request all authors and reviewers to inform us about any potential conflicts of interest (e.g., financial, personal, political, or academic) that could affect their judgment. The corresponding author must sign a *Conflict of Interest Form* indicating that all authors have significantly contributed to the research and meet the authorship criteria.

Table of Contents

English Teachers' Perception of School Climate and Effective Teaching

Karim Sadeghi[1,*] iD & Fatemeh Esmaeili[2]

[1] Urmia University, Urmia, Iran
[2] PhD Graduate, Urmia University, Urmia, Iran

Abstract
The present study investigated Iranian English as a Foreign Language (EFL) teachers' perception of school climate and its relationship with their conception of effective teaching. A body of 147 teachers selected through purposive and snow-ball sampling strategies answered Oderand and Eisenschmidt's (2016) questionnaire about school climate and effective teaching. Data were analyzed using factor analysis, which produced three factors to explain teachers' perception of school contextual factors and three for their conception of effective teaching. The results indicated that teachers' perception of school contextual factors to some degree correlated with their approach to teaching in the classroom. The study highlights the fundamental role of school senior management in helping teachers to apply appropriate teaching approaches.

1. INTRODUCTION

the nature of educational system is experiencing a change in the 21st century. A global aim for governments and educators is to find the most effective ways of teaching in order to improve learning. The curricula are not seen as meeting the needs of students within a globalized world (Hope, 2015). Individual learners are in the center of education and their needs, abilities, skills, interests and individual characteristics are accounted for by curriculum designers and policy makers (Wenden, 2002). Students need to develop different competences and skills such as creativity, problem solving, critical thinking, and interpersonal skills in order to cope with the demands of a highly advanced society (Jerald, 2009).

In order to accommodate students' needs in education in the current century a Fundamental Reform Document of Education (FRDE) was proposed in Iran in 2011. The FRDE emphasized enhanced application of active, creative and enlightening approaches, utilization of modern educational and training equipment and technologies in conformity with the objectives of education, highlighting individual differences among students and promotion of teachers' professional training system in the Ministry of Education with an emphasis on sustainable interaction of teachers with schools and scientific–research institutes during their studies and provision of opportunities to gain useful experience from classrooms and educational spaces.

---

* Corresponding Author: Karim Sadeghi, Department of English Language and Literature, Urmia University, Urmia, Iran. **Email**: k.sadeghi@urmia.ac.ir

One important skill for students in todays' advanced world is learning English language. English as an international language possesses a high status in Iranian society and plays a critical role in higher education. It is indeed viewed as a sign of educational and social achievement (Sadeghi & Richards, 2015). English is a vehicle for educational advancement (Farhady, Sajadi Hezaveh, & Hedayati, 2010) and is a gate for new knowledge and technology. The rapid changes in the Iranian society and the approach to teaching English have transformed it into a fashionable trend which has resulted in the spread of English language within the country (Davari, 2013).

Over the past years, the Ministry of education has attempted to make changes in approaches to teaching English in schools. In 2003, Iranian pre-university English syllabus underwent a reform. The principal goal was to propose a shift away from the long-established grammar-translation curriculum practice towards teaching for communicative competence. In 2007, the first Iranian national curriculum for teaching foreign languages (based on communicative approach teaching) was developed by a team working under the supervision of the Ministry of education. In 2011, the new curriculum was implemented and the books and evaluation system underwent radical changes, which came through as a result of changes in policy toward teaching English in Iran.

Application of FRDE in line with the new curriculum in English language teaching at schools requires an emphasis on individual learners' needs and interests and following learner-centered teaching. In the application of learner-centered teaching an important role is played by teachers, because teachers are ultimate agents of educational change and the classroom is the main place for innovation or an obstacle to it; accordingly, no plan for a long- lasting educational change can ignore teachers' roles (Hargeaves & Shirly, 2009). It is clear that teachers' epistemological beliefs shape their teaching practices in the classroom) Ethell, 1997; Kogan, 1992; Woolfolk Hoy, Hoy, & Davis, 2009; Rubie, Davis, Flint, & McDonald, 2012).

With the current interest in developing students' functional skills such as learning English language and applying learner- centered teaching, the educational systems at school level has been more seriously challenged. The main perspective is that teachers alter their instructional practices according to school contextual variables (Solomon, Battistich, & Hom, 1996). Schools are the main site for professional learning. Schools are the place in which teachers are provided with opportunities to cooperate with their colleagues and construct knowledge in authentic contexts (Maloney & Konza, 2011).

The application of principles of FRDE and communicative language teaching with a focus on student needs is affected by school climate, which plays a critical role in learner success and achievement of pedagogical goals. Thus, it is critical to find out teachers' perceptions of school contextual factors and whether school climate helps them in achieving their teaching aims. In addition, utilizing learner-centered education depends on teachers' belief about appropriate ways of teaching and what they consider as fundamental components of teaching process. Exploring the relationship between teachers' perceptions of school climate and their conception of effective teaching can help in removing obstacles in teaching and attaining the goals of education in the current century.

2. LITERATURE REVIEW

School climate

Although there are wide definitions of school climate and different aspects of it are emphasized, the fundamental view in all of them is the existence of positive social relationships and interpersonal interactions (Hayness, Emons, & Ben-Avie, 1997). Cohen, Macabe, Michelli, and Picherall (2009) defined it as the quality of school life which encompasses the norms, objectives, values, interpersonal relationships, learning atmosphere, and organizational structure of

school and is connected with feeling socially, emotionally, and psychologically safe. School climate is an important area of inquiry since it is associated with students' academic, social and psychological outcomes (Anderson, 1982; Koth, Bradshaw, & Leaf, 2008). Emmons, Comer and Hayques (1996) defined school climate as shared beliefs and values that shape interactions between students, teachers, and administrators. Positive school climate is identified as a fundamental aspect of successful schools (Brands, Felner, Shim, Sseitsinger, & Dumas, 2003). Schools are expected to cater for individual student's needs and their conceptual development (Coofey, 2007). Educators and researchers have acknowledged that a positive climate plays a fundamental role in high academic achievement (Thepa, Cohen, Guffey, & D'Allesandro, 2013).

Teachers' perception of school climate is related to their job satisfaction (Taylor & Tashakkori, 1995), burnout and attrition (Perie, & Baker, 1997; Grayson & Alvarez, 2008; Keiger, 2010), higher teacher commitment (Hoy, 1990), school cohesiveness and sense of trust (Bryk & Schneider, 2002), cooperation in school community (Lunenburg, 2010), and their desire to apply new academic principles and curricula (Beets, Flay, Vuckinich, Acock, Li, & Allred, 2008). A complex range of internal and external factors such as interpersonal experiences with student and school personnel shape individual and collective experiences of teachers. It is necessary that school principals and administrators support the community members' needs and views and provide a situation in which the members including teachers' work collaboratively to achieve academic objectives. Teachers should share community leadership and feel ownership of work for it (Boone, 2010). In fact, research indicates that teacher authority and their chance of being involved in decision-making process at school culminates in positive changes in their profession. Vangrieken, Meredith, Packer, and Kyndt (2017) indicated that school leadership exerts a significant effect on teacher communities.

Conception of effective teaching

Teachers' conceptions of effective teaching and the efficiency of different methods and approaches have been the subject of many research studies (Gao & Watkins, 2002; Kember, 1997; Lam & Kember, 2004; Pratt, Arseneau, & Collins, 2001; Samuelowicz & Bain, 2001). Although there are different terminologies to refer to the best way of teaching, such as 'excellent teaching' (Chen, 2007; Chen et al., 2012; Kane, Sandretto, & Heath, 2004), 'highly accomplished' (National Board for Professional Teaching Standards, Serafini, 2002), 'effective' (Aregbeyen, 2010; Arikan, Taser, & Sarac-Suzer, 2008; Ganjabi, 2011), 'better teaching' (Gore, Griffiths, & Ladwig, 2004), and 'good teaching' (Kember & Kwan, 2000; Sakurai, 2012), all of them focus on similar aspects of teaching practice. Kember (1997) in his framework for describing teacher conception identified three main factors: "teacher-centered" orientation which deals with transmitting information and structured knowledge, "student-centered" orientation which relates to facilitating knowledge and encouraging conceptual changes and "student- teacher" orientation which aims to combine the two orientations and reduce the controversies relating to them.

Studies of teacher conception of effective teaching lie within two main paradigms (Lopez-I~niguez & Poz, 2014). One emphasizes transmitting established knowledge and the other focuses on developing students' higher order thinking and capabilities by applying constructive views of teaching (Winne & Nesbit, 2010). A new position is recently taken by educators which emphasizes teachers' responsibility for transmitting or conveying knowledge, but also takes into account a learner's active role in learning process (Martın, Pozo, Mateos, Martın, & Perez-Echeverrıa, 2012, cited in Lopez- I~niguez & Poz, 2014). Although the transition model of teaching is the dominant approach to compulsory education in much of the world (OECD, 2009), many have criticized it for being ineffective. Developing higher order thinking skills such as problem solving, critical

thinking and creativity has instead become the essential components of the curriculum in the 21st century (Lam & Lipstone, 2001; Saavedra & Opfer, 2012). Excellent teachers put emphasis on the nature of task because it increases student engagement and helps developing higher level thinking skills and concepts (Boston & Smith, 2009). Studies show that transmission discourages students from applying deep approaches to learning (Gow & Kember, 1993) and improvement in education is achieved by teachers' changing their views of teaching and learning (Gibbs, 1995). Effective teaching is the basis for achieving a wide range of goals in society, including enhancing student achievement, improving school conditions and organizing appropriate teacher education programs (Betore & Artiga, 2004, cited in Chen, 2012).

Chen (2007) studied middle school teachers' conceptions of effective teaching on the north of China. The study confirmed that four main principles were identified as important in teachers' views: (1) caring for students, (2) guiding students' overall development, (3) connecting school knowledge to other areas, and (4) planning structured lessons.

Bear, Gaskins, Blank, and Chen (2011) studied students' perception of school climate and found five main factors as critical: 1) teacher- student relations, 2) student-student relations, 3) fairness of values, 4) liking of schools, and 5) school safety.

Tavakoli and Baniasad-Azad (2016) studied Iranian high school teachers' conceptions of effective teaching and the relationship of those conceptions to their teaching practices. The results indicated that Iranian teachers identified "focus on students" and "being exam- oriented "as main features of effective teaching.

Review of related literature indicates that in spite of the emphasis on learner-centered and communicative teaching in Iran, it is not clear how schools shape teachers' orientation to teaching task. There is lack of research on the role of school contextual factors and teachers' conception of effective teaching. Therefore, to fill this gap, the present study aims to study Iranian EFL teachers' perceptions of school climate and find its relationship with their conception of effective teaching. The research questions addressed are:

1. What are Iranian EFL teachers' perceptions of their school climate?
2. Is there any significant relationship between Iranian EFL teachers' perceptions of school climate and effective teaching?

3. METHOD

Participants

A cohort of 147 Iranian EFL teachers teaching at public school participated in this study. Their age range was 25-61 and they were chosen from two provinces: East Azerbaijan (70 %) and West Azerbaijan (30%). The experience of participants ranged from 2 to 30 years and the average year of teaching experience was 15 years. Sixty percent of the participants had a BA degree and 33% percent had a Master's degree. Ph.D. candidates accounted for 7% of population.

Design

The current study is a kind of survey study, which aims to explore Iranian EFL teachers' perception of school climate and its relationship with approaches to teaching at the classroom.

The data were collected through purposive convenience and snowball sampling strategies. We took questionnaire to schools which we had access to and sent it to telegram groups of EFL teachers. Also, we asked respondents to introduce their colleagues to us. In the process of data analysis, we conducted Factor analysis, which aims to identify underlying variables, or factors that explain the pattern of correlations within a set of observed variables. Factor analysis is often used in data

reduction to identify a small number of factors that explain most of the variance observed in a much larger number of manifest variables.

Instruments

To gather data for this survey, we used the questionnaire developed by Oder and Eisenschmidt (2016). This questionnaire deals with teachers' perceptions of school climate and effective teaching. The basis for this questionnaire is principals of OECD Teaching and Learning International Survey (TALES) questionnaire (OECD, 2009). The questionnaire consists of 29 items; section one relates to teachers' perception of school climate involving 13 items, and section two considers teachers' conception of effective teaching and includes 16 items. The data were analyzed using SPSS Statistical Package for the Social Sciences version 24. The mean scores with standard deviation were calculated for each item and Principal Components Analysis using varimax rotation with Kaiser Normalization was used for conducting factor analysis for the two sections of questionnaire.

Procedure

The data for present study was collected during 2016-2017 school year. In the first phase of data collection, we took the questionnaires to schools and asked teachers to complete them. It took them about ten minutes to answer the questionnaire. In second phase of data collection, we shared the questionnaire with Language Teacher Groups in telegram and asked for teachers' cooperation. About 30% of data was collected in this way. And finally using snowball sampling, we asked our participants to recruit other subjects. Some of them introduced their friends or colleagues and we sent the file via email or telegram to these teachers. We gathered about 40% of our data in this way.

Data analysis

In the process of data analysis, we conducted factor analysis in order to detect underlying variables, and to find the correlation between factors emerging from school climate and effective teaching, we applied Pearson r correlation coefficient to data.

As mentioned above, the module of school climate (module A) consisted of thirteen items. In order to reduce it to manageable factors we used rotated component analysis. The communalities of items were checked and items with low values were removed. The results of factorization indicated that seven items, which accounted for 65% of variance items produced three main factors describing teachers' perception of school climate in Iran.

The schools were characterized by supportive leadership (factor 1), which was reflected in senior managements' readiness to advocate and support teachers whenever there was a problem, teachers' active involvement in school decision-making and value attached to teachers' good work. The second factor emerging from data was related to boring and frustrating depiction of teaching experience in which the task teachers faced was routine and there was little place for innovation and variety in teaching. Second factor also indicated teachers' lack of eagerness to their job and their preference for giving up their job. The third factor accounted for cooperative climate in which there was a strong relationship among colleagues at schools. Teachers perceived their school climate as conducive for cooperative work in which teachers tried to experience mutual understanding and work. Internal consistency of factors was checked with Cronbach's alpha coefficients, which produced following results F1: inclusive leadership. $\alpha = 0.72$, F2: frustrating teaching experience $\alpha = 0.70$, F3: cooperative climate: $\alpha = 0.66$.

The factors we obtained from present data are to some extent similar to what Oder and Eisenschmidt (2016) found in their study, only one factor was different for the two studies. Their study produced three main factors related to teachers' perception of school climate in Estonia.

Inclusive leadership (factor 1), inspiring climate (factor 2) and cooperative climate (factor3) described Estonian EFL teachers' perception of school climate.

The module of effective teaching (Module B) comprises 16 items. Replicating the design of the questionnaire, we divided the items of effective teaching into two groups: (a) traditional and (b) constructivist teaching. Traditional teaching comprises four items ($\alpha = 0.62$) emphasizing teacher-centered banking approach to education. The constructivist teaching that consists of 12 items ($\alpha = 0.77$) advocates learners' active role in the classroom and learning process, discovery and inquiry approaches to teaching, and importance of group work in constructing new knowledge.

The process of item factorization for constructive approach to teaching produced three main factors. Integrative teaching involved integration of different subject areas and learners' active role in learning process (factor1). Active learning involves providing students with opportunities for constructing new knowledge and doing practical tasks. Authentic and meaningful teaching which involves using real life every day experiences in language teaching. These three factors accounted for 66% of total variance. Internal consistency of factors checked with Cronbach's alpha coefficients produced following results: Integrative teaching (F1) $\alpha = 0.70$, active learning (F2) $\alpha = 0.70$ and authentic and meaningful teaching (F3) $\alpha = 0.66$. (see Table 1). Oder and Eisenschmidt (2016) also found three main factors to account for constructive teaching: Learner independence; Active learning; Integrative practice.

4. RESULTS

Bellow we first provide descriptive statistics for Iranian EFL teachers' perceptions of school climate and their conception of effective teaching (Tables 2 and 3). Then the relationship between teachers' perception of school and effective teaching are presented. As Table 2 shows, item 1 and item 4 have the highest mean, and items 3, and 13 have the lowest mean.

Table 1: Sample Items and Factor Reliabilities (Cronbach's alpha coefficients)

	Factors	Cronbach's alpha	Sample items
School Climate	Supportive administration	a=0.72	I know that school administration is ready to help me
	Frustrating experience	a=0.70	I prefer someone else take responsibility
	Cooperative climate	a=0.66	I often cooperate with my colleagues
Effective Teaching	Integrative teaching	a=0.70	Knowledge and skills of other subjects should be integrated
	Active learning	a=0.70	Teachers should guide students to discover and construct new knowledge
	Authentic &meaningful teaching	a=0.66	Real problems and future life make a meaningful context for students' knowledge development

Table 2: Descriptive Statistics for English Teachers' Receptions of School Climate

	Mean	SD	N
A 1. I often cooperate with my colleagues	4.28	.738	147
A 2. Sometimes my colleagues do not cooperate with me	3.27	.922	147
A 3. The physical environment at our school is unpleasant	3.12	1.12	147
A 4. The necessary teaching resources (course-books, additional materials, copying machine, library, media resources) are available at school	3.37	1.11	147
A 5. Teachers can greatly influence major decisions connected with school and creation of school policy	3.21	1.34	147
A 6. I prefer someone else to take responsibility	3.74	1.13	147
A 7. Teachers' work consists of routine actions	3.12	1.07	147
A 8. Teaching inspires me to be creative	4.16	.836	147
A 9. Teachers' good work is recognized at our school	3.50	1.10	147
A 10. Nobody tells me I am a good teacher	3.47	1.07	147
A 11. I know that school administration is ready to help me with problems arising in class	3.49	1.15	147
A 12. I look forward to work every day	3.62	1.07	147
A 13. My health problems could be connected with stress derived from teaching	2.99	1.18	147

Table 3: Descriptive Statistics for English Teachers' Conception of Effective Teaching

	Mean	SD	N
B1 Real problems and future life make a meaningful context for students' knowledge development	4.12	.928	147
B2 Teaching should be based on questions with clear and correct answers and ideas that most of the students can grasp quickly	2.61	1.28	147
B3 How much students learn depends on their background knowledge, which is why teaching facts are so important	2.70	1.15	147
B4 Effective teachers show their students how to solve problems and tasks	4.12	.945	147
B5 My role as a teacher is to support students' problem-solving activities	4.19	.888	147
B6 Students learn best solving problems independently	3.01	1.20	147
B7 Students should be given possibilities to try out practical tasks themselves before teacher's explanation	3.83	.891	147
B8 Teachers should guide students to discover and construct new knowledge	4.12	.870	147
B9 Teachers should vary their teaching methods to achieve students' meaningful learning and understanding	4.17	.825	147
B10 Students should have possibilities for discussing new ideas and listen to their peers' opinions in small groups	4.38	.686	147
B11 Thinking and discussion are more important than curricular subject requirements	3.98	2.11	147
B12 Most student activities in class require previous knowledge and skills being used in a novel way	3.48	.878	147
B13 Knowledge and skills of other subjects should be integrated into teaching	3.42	1.04	147
B14 Students and teachers create assessment forms and criteria together	3.43	1.03	147
B15 Assessment should include practical tasks, projects and student research	3.74	J.927	147
B16 In general, effective learning requires silence in the class	3.15	1.35	147

Table 4: Correlation between Teachers' Perception of School Climate and Approaches to Teaching

		Integrative Teaching	Active Teaching	Authentic Teaching	Traditional Teaching
Supportive Administration	Pearson correlation Sig.(2tailed) N	.430 .000** 147	.055 .517 147	-.033 .693 147	-.172 .036** 147
Frustrating Teaching Experience	Pearson correlation Sig.(2tailed) N	-.147 .057 147	-.019 .823 147	-.032 .704 147	.171 038** 147
Cooperative Climate	Pearson correlation Sig.(2tailed) N	.111 .180 147	.130 .117 147	.044 598 147	.096 .245 147

** Correlation is significant at the 0.05 level (2-tailed).

Descriptive statistics for Iranian EFL teachers' conception of effective teaching are given in Table 3. The results show that item 10 (Students should have possibilities for discussing new ideas and listen to their peers' opinions) has the highest mean followed by item 9 and item 8. On the other hand, items 2 and 3 have the lowest mean of effective teaching.

Table 4 shows correlation results between school climate factors and approaches to teaching. Teachers' perception of school climate characterized as supportive administration (factor 1) has a significantly positive correlation with effective teaching characterized as integrative approach but it has a significantly negative correlation with traditional approaches to teaching. However, supportive leadership does not have a statistically significant correlation with active and authentic elements of effective teaching.

Teachers' perception of school climate as being frustrating (factor 2) has a statistically significant correlation with traditional approach to teaching. However, it does not have any significant correlation with factors of effective teaching. Teachers' perception of school climate as being a cooperative setting (factor 3) does not have statistically significant correlation with approaches to teaching.

5. DISCUSSION

The present study investigated the relationship between EFL teachers' perception of school climate and their conception of effective teaching. The results indicated that supportive administration was the first factor to describe school climate in Iran. And this highlights the critical and fundamental role of school principals in schools. As Halling (2003) notes principals are vital in school since they play managerial, political, institutional and instructional roles. Various studies have indicated direct and significant roles of managers in school success, effective teaching, and learning process and outcome (Hallinger & Heck 1998; Harris, 2009; OECD, 2001; Robinson, Lloyd, & Rowe, 2008). School leaders play a critical role in ensuring success and have the responsibility to provide a safe and challenging environment in which teachers can experience collective inquiry, and make sure that the necessary resources are available for teachers.

The results of the current study indicated that supportive administration as one of the factors of teachers' perception of school climate correlated negatively with traditional teacher-oriented teaching. The result reflects the importance of senior management's support in shaping teachers' orientation to teaching task. This is partially in contrast with Oder and Eisenschmidt's (2016) finding that traditional approach to teaching did not have any correlation with any school factor.

However, Oder and Eisenschmidt (2016) also found that inclusive leadership correlated positively with learner-centered teaching. The results are also in line with the finding by Vangrieken, Meredith, Packer, and Kyndt (2017), which indicated that school leadership exerts a significant effect on teacher communities. Similarly, Allensworth, Ponisciak, and Mazzeo (2009) confirmed that the level of support provided by school management and teachers' active involvement in school decision making process affects teachers' approaches to teaching.

Whenever teachers are advocated in a school setting and the manager is ready to help them in difficult situations, they try to avoid traditional transmitted or banking approach to teaching and seek alternative approaches such as learner-centered and constructive teaching. The reason for this might be that students and teachers are accustomed to teacher-oriented classes in which students as passive agents act under the authority of the teacher. In fact, a strict discipline and an absolute teacher authority are distinguishing characteristics of Asian teaching contexts (Tan, 2006). Any switches to constructive teaching, transition in role relationships and authority sharing between teachers and students may cause controversies and crises in teaching process since neither teachers nor students are familiar with principles and rules of learner-centered education. Teachers also need more authority in order to make changes in established curriculum and syllabus for the course. Any attempt to apply constructive teaching in which students' individuality, needs and independency are emphasized would affect different aspects of teaching profession such as syllabus, classroom disciplinary climate, collaborative effort, professional development and this requires the support of senior management at school level and the government and ministry of education at higher levels.

The most widely practiced method in foreign language teaching is grammar translation method, which is relatively easy to apply, and makes few demands on teachers (Qing-xue & Jin-fang, 2007). Any shift from traditional approaches to constructive i.e. communicative approach to language teaching is new attempt, which requires highly professional teachers. However, as OCED (2009) reports most of the schools in most countries lack professional teachers. Teachers' lack of professional development and demanding pressure to develop the required skills to apply constructive teaching may be another reason for the importance of supportive administration in schools. Because teachers' self-esteem in their job is closely linked to a sense of professional efficacy (Nias, 1996), supportive administrators should help teachers improve their professional skills by providing opportunities for them to take part in teacher training courses, workshops and conferences.

School administrators promote teaching by not only creating organizational conditions but also by providing opportunities for planned discussion and mutual reflection. Results of the current study indicated that teachers believed a supportive administrator would inspire them to apply integrative teaching. In order to follow an integrative curriculum at school level which aims to encourage students to see the interconnectedness and interrelationship between the curriculum areas and which views students as active learners capable of interpretation and communication, there is need for a coherent and systematic plan and organization at school which brings the whole community of practice together. School administrators can help teachers follow an integrative approach by actively engaging the community to create shared responsibility for student success in different subject areas. Integration of different subject areas with an aim to enhance student motivation and creativity and empower them by providing meaningful learning requires teachers' acceptable level of knowledge in different subject areas and this surely necessitates teachers' collaborative work and their authority to be involved in school related decisions. School principals and teachers need to come to a mutual understanding that integrative teaching not only results in

increased student learning but also helps teachers grow and reflect in areas that they do not have higher expertise. Furthermore, this can be a great achievement and a long-term investment for teachers and schools.

The second factor that teachers perceived to describe school climate was frustrating and boring view of teaching which is linked to the notion of teacher burnout and attrition. Teacher attrition is a widespread problem all over the world which affects teachers' performance and their motivation to teach, and as Chan (2003) states "it might impair the quality of teaching as well as leading to job dissatisfaction, work alienation, physical and emotional ill-health and teachers leaving the profession' (p. 382). Teacher burnout can be caused by personal as well as organizational factors such as poor salary, student discipline problems, poor administrative support, poor student motivation and lack of faculty influence (De Heus & Diekstra, 1999; Ingersoll, 2001; Ingersoll & Smith, 2003; Leiter & Masalch, 2005; Sadeghi & khezrlu, 2016).

The results of the current survey indicated that EFL teachers perceived their teaching experience frustrating in spite of the supportive administration and strong collegial relationship among school members. The reason for this might be that although, senior management and cooperation among teachers are fundamental to determining the level of teacher accomplishment and job satisfaction, they are not the sole cause of it; various factors come into play. This can be related to challenges of teaching English at schools. As Loh and Liew (2016) reported English teachers are confronted with problems such as sociocultural values attached to teaching English, the tension of grading student essays, the performance pressures of high-stakes testing, and the need for culturally responsive pedagogies. Naylor and Malcomson (2001) also reported that Canadian English teachers found grading a stressful task in their occupation, and they were forced to reduce the number of their classes (cited in Loh & Liew, 2016). Another problem for teachers might be the constraints and institutional rules of teaching appropriately. Teachers' attempt to teach based on their own values and applying innovative methods is constrained by common belief in teacher-oriented pedagogies and the tension of national high-stakes examinations (Curdt-Christiansen & Silver, 2013). Teachers have to follow what is known as "'teach to the test'… with less time devoted to activities that are not part of the test" (p. 786). Education policies, which advocate communicative teaching are not consistent with the demands of national examinations (Littlewood, 2007). Teacher' abandoning of their values in teaching and not being able to apply what they believe because of external pressures and constraints within their schools is closely linked to a sense of loss and frustration on them (Hargreaves, 2003; Nias, 1996).

Feeling of frustration and exhaustion that teachers reported in their perception of school climate correlated with traditional teaching approaches. The results confirm findings of previous studies which indicated that teacher burnout can be closely linked to lower achievement levels since burnout teachers do not exert the effort once they did and try to reduce their relationship with students (Farber, 1991). When teachers are emotionally exhausted for any reason they opt for traditional and less effective teaching approaches because they feel that they cannot accomplish their goals. This will have a severe effect on learning outcome in language classes. Applying learner-centered approaches such as communicative language teaching requires teacher creativity, flexibility, professional development,collaboration and communication skills all of which require teacher motivation and patience which burnout teachers do not demonstrate in their work.

The third factor which teachers perceived to describe school climate was cooperation and strong collegial relationship among teachers. Teachers reported that they attempted to work cooperatively with their colleagues. School- based teacher collaboration stirs up greater improvements in teaching and learning and brings about changes in teachers' ideology and their approaches to

teaching (Clement & Vandenberghe, 2000; Fernandez & Yoshida, 2004). A positive school climate characterized by strong collaborative communities plays a critical role in teachers' performance (Talbert & Milbrey, 2002). It also provides possibilities for new models of professional development based on shared reflection in the workplace (Lieberman, 1996).

An unexpected result that we obtained was that teachers' perception of school climate, as a cooperative site did not correlate with any of components of constructive teaching. This is in contrast with results obtained by Oder and Eisenschmidt (2016). In their study cooperative school climate correlated with effective teaching. The results of the study might confirm Huberman's position that teachers seek isolation after tenure and that school community and collaboration might not lead to lasting instructional changes (1993). He argued against benefits of collaboration among colleagues and believed that school bureaucratic mandates which aim to improve collaboration at school undermines rather support teachers' instruction and effectiveness (Talbert & Milbrey, 2002).

One reason for the result we obtained might be difficulties involved in managing classroom events. Maybe the realities of classroom or the facilities do not allow for utilization of learner-centered approaches in spite of shared understanding and cooperation among teachers. Another reason might be that teachers overestimated cooperation with their colleagues and it is not clear what their cooperative work involves.

6. CONCLUSION

The results of the current study accentuate the importance of school senior management in paving the way for EFL teachers' effective teaching in the classroom. Teachers need leadership support in order to move from traditional to constructive teaching approaches and this is linked to different aspects of leadership role in schools. A good and well- supported leadership in schools can create a sense of ownership and purpose in teachers' approach to their work. Improving teachers' professional autonomy makes teaching an attractive profession and improves the quality of the classroom teaching practice (OCED, 2001). A very important implication of this study relates to the need for improving school leadership as major school reform agenda. School improvement movements in the past 20 years have put a great emphasis on the role of school managers. Government and policy makers should seek, formulate and implement school leadership policies and programs in order to improve education within the country since efficient education system is critical to country's advancement. School managers' roles, competencies and responsibilities should be redefined and broadened.

The results of this study partially reveals that changes in approaches and textbooks cannot guarantee effective approaches by teachers; still teachers find their task as a routine and boring and this severely affects student outcome and can lead to students' reduced motivation. OCED (2001) reports that in 20 out of 28 countries, more than one in four 15-year- old students view school as a place that they would not like to go and in almost half the countries students find school boring and frustrating. Due to the critical role played by teachers, it is vital for teachers to regain their energy, motivation and do their job eagerly. Identifying teachers' needs and providing effective support at school and classroom level can contribute to enhanced motivation. In order to make changes the government, policy makers and school managers should bring variety to school activities and provide teachers with a chance to experience innovative and creative teaching process. This can be achievable by decreasing the number of students in classrooms, bringing technology into classrooms and providing more opportunities for teachers to take part in conferences, workshops and other professional development activities. Involving teachers in different work-related practical projects could also change their prospective.

The present study has some limitations. First, there is a need for exploratory qualitative study on EFL teacher's perception of school climate and effective teaching. Applying mixed methods could give a clearer picture of EFL teachers' views, skills, preferences and obstacles in their teaching approaches. Second, the data for the present study was collected from public schools without considering economic and social issues. Involving private school teachers' perceptions can give illuminating insights since the culture, social class, facilities, and education policies are totally different in these two types of schools.

Teachers' perception of school climate can be affected by their emotions, personal and professional identities. Future research can delve into the issue by considering interconnection between social, cultural and political complexities of schools, teachers' professional identity and their teaching practices. Future research can also be replicated using school senior managements' perceptions of school contextual factors and finding the extent to which their authority and instructional policies can help improve the quality of language teaching at schools.

References

Allensworth, E., Ponisciak, S., & Mazzeo, C. (2009). *The schools teachers leave: Teacher mobility in Chicago Public Schools*. University of Chicago.

Anderson, C. S. (1982). The search for school climate: A review of the research. *Review of Educational Research, 52*(3), 368-420. https://doi.org/10.3102/00346543052003368

Aregbeyen, O. (2010). Students' perception of effective teaching and effective lecturers' characteristics at the University of Ibadan, Nigeria. *Pakistan Journal of Social Sciences, 7*(2), 62–69.

Arikan, A., Taser, D., & Sarac-Suzer, H. S. (2008). The effective English language teacher from the perspectives of Turkish preparatory school students. *Education and Science, 33*, 42–51.

Bear, G. G., Gaskins, C., Blank, J., & Chen, F. F. (2011). Delaware School Climate Survey– student: Its factor structure, concurrent validity, and reliability. *Journal of School Psychology, 49*(2), 157–174. https://doi.org/10.1016/j.jsp.2011.01.001.

Beets, M. W., Flay, B. R., Vuckinich, S., Acock, A. C., Li, K. K., & Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the Positive Action Program: A diffusion of innovations model. *Prevention Science, 9*(4), 264–275. https://doi.org/10.1007/s11121-008-0100-2

Boon, S. C. (2010). Professional learning communities' impact: A case study investigating teachers' perceptions and professional learning satisfaction at one urban middle school (Doctoral dissertation). ProQuest Dissertation and Theses. UMI NO. 3419559.

Boston, M., & Smith, M. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119-156.

Brand, S., Felner, R. D., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, cultural pluralism and school safety. *Journal of Educational Psychology, 95*(3), 570–588. https://doi.org/10.1037/0022-0663.95.3.570.

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for school improvement*. Russell Sage Foundation.

Chan, D. W. (2003). Hardiness and its role in the stress-burnout relationship among prospective Chinese teachers in Hong Kong. *Teaching and Teacher Education, 19*(4), 381-395. https://doi.org/10.1016/S0742-051X(03)00023-4(

Chen, J. (2007). Teacher's conceptions of excellent teaching in middle school in the North of China. *Asia Pacific Education Review, 8*(2), 288–297. https://doi.org/10.1007/BF03026464

Chen, J., Brown, G., Hattie, J., & Milward, P. (2012). Teachers' conceptions of excellent teaching and its relationships to self-reported teaching practices. *Teaching and Teacher Education, 28*(7), 936-947. https://doi.org/10.1016/j.tate.2012.04.006

Clement, M., & Vandenberghe, R. (2000). Teachers' professional development: A solitary or collegial adventure? *Teaching and Teacher Education, 16*(1), 81–101. https://doi.org/10.1016/S0742-051X(99)00051-7

Cohen, J., McCabe, E. M., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, teacher education and practice. *Teachers College Record, 111*(1), 180-213.

Coffey, S. (2007). Differentiation in theory and practice. In J. Dillon & M. Maguire (Eds.), *Becoming a teacher: Issues in secondary education* (3rd ed., pp. 187–201). Open University Press.

Curdt-Christiansen, X. L., & Silver, R. (2013). New wine into old skins: The enactment of literacy policy in Singapore. *Language and Education, 27*(3), 246-260. https://doi.org/10.1080/09500782.2012.669048

Davari, H. (2013). *English language teaching and linguistic imperialism*. LAP Lambert Academic Publishing.

DeHeus, P., & Diekstra, R. F. W. (1999). Do teachers burnout more easily? A comparison of teachers with other social professions on work stress and burnout symptoms. In R. Vandenberghe & A. M. Huberman (Eds.), *Understanding and preventing teacher burnout: A source book on international research and practice* (pp. 269-284). Cambridge University Press.

Emmons, C. L., Comer, J. P., & Haynes, N. M. (1996). Translating theory into practice: Comer's theory of school reform. In J. P. Comer, N. M. Haynes, E. Joyner, & M. Ben-Avie (Eds.), *Rallying the whole village* (pp. 27–41). Teachers College Press.

Ethell, R. G. (1997). Reconciling propositional and procedural knowledge: Beginning teachers' knowledge in action (Unpublished doctoral thesis). Griffith University, Brisbane, Australia.

Farber, B. A. (1991). *Crisis in education: Stress and burnout in the American teacher*. Jossey Bass.

Farhady, H., Sajadi Hezaveh, F., & Hedayati, H. (2010). Reflections on foreign language education in Iran. *The Electronic Journal for English as a Second Language, 13*(4), 1-18.

Fernandez, C., & Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning*. Lawrence Erlbaum Associates.

Ganjabi, M. (2011). Effective foreign language teaching: A matter of Iranian students' and teachers' beliefs. *English Language Teaching, 4*(2), 46–54. https://doi.org/10.5539/elt.v4n2p46

Gao, L., & Watkins, A. D. (2002). Conceptions of teaching held by school science teachers in R. China: Identification and cross-cultural comparisons. *International Journal of Science Education, 24*(1), 61–79. https://doi.org/10.1080/09500690110066930

Gibbs, G. (1995). Changing lecturers' conceptions of teaching and learning through action research. In A. Brew (Ed.), *Directions in staff development* (pp. 21-35). Society for Research into Higher Education and Open University Press.

Gore, J., Griffiths, T., & Ladwig, J. (2004). Towards better teaching: Productive pedagogy as a framework for teacher education. *Teaching and Teacher Education, 20*(4), 375–387. https://doi.org/10.1016/j.tate.2004.02.010

Gow, L., & Kember, D. (1993). Conceptions of teaching and their relationship to student learning. *British Journal of Educational Psychology, 63*(1), 20-33. https://doi.org/10.1111/j.2044-8279.1993.tb01039.x

Grayson, J., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education, 24*(5), 1349-1363. https://doi.org/10.1016/j.tate.2007.06.007

H, D., Lee, M., & Teng, Y. (2016). Exploring the relationship between school-level teacher qualifications and teachers' perceptions of school-based professional learning community practices. *Teaching and Teacher Education, 54*, 32-43. https://doi.org/10.1016/j.tate.2016.01.002

Halling, M. A. (2003). Stress and stress management strategies among elementary school principals (Doctoral dissertation, Florida Atlantic University). Retrieved from http://proquest.umi.com.ezproxy.fau.edu

Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. *School Effectiveness and School Improvement, 9*(2), 157-191. https://doi.org/10.1080/0924345980090203

Hargreaves, A. (2003). *Teaching in the knowledge society: Education in the age of insecurity*. Teachers College Press.

Hargreaves, A., & Shirley, D. (2009). *The fourth way: The inspiring future for educational change*. Corwin Press.

Harris, A. (2009). *Distributed leadership*. Springer.

Haynes, N. M., Emmons, C., & Ben-Avie, M. (1997). School climate as a factor in student adjustment and achievement. *Journal of Educational and Psychological Consultation, 8*(3), 321-329. https://doi.org/10.1207/s1532768xjepc0803_7.

Huberman, M. (1993). The model of the independent artisan in teachers' professional relations. In J. W. Little & M. W. McLauglin (Eds.), *Teachers' work: Individuals, colleagues, and contexts* (pp. 1–19). Teachers College Press.

Henry, G. T., Kershaw, D. C., Zulli, R. A., & Smith, A. A. (2012). Incorporating teacher effectiveness into teacher preparation program evaluation. *Journal of Teacher Education, 64*(5), 335–355. https://doi.org/10.1177/0022487113492300

Hope, M. (2015). Conceptions of person-centered education in the 21st century: Developing a position statement for ways of being in educational organizations. *Experiential Psychotherapies, 13*(3), 217–232. https://doi.org/10.1080/14779757.2014.969294

Hoy, W. K. (1990). Organizational climate and culture: A conceptual analysis of the school workplace. *Journal of Educational and Psychological Consultation, 1*(2), 149-168. https://doi.org/10.1207/s1532768xjepc0102_4

Ingersoll, R. (2001). *Teacher turnover, teacher shortages and the organization of schools*. Center for the Study of Teaching and Policy, University of Washington.

Ingersoll, R. M. (2003). *Is there really a teacher shortage?* Center for the Study of Teaching and Policy, University of Washington.

Jerald, C. D. (2009). *Defining a 21st century education*. Centre for Public Education.

Kane, R., Sandretto, S., & Heath, C. (2004). An investigation into excellent tertiary teaching: Emphasizing reflective practice. *Higher Education, 47*, 283–310. https://doi.org/10.1023/B:HIGH.0000016442.55338.24

Keigher, A. (2010). *Teacher attrition and mobility: Results from the 2008-09 Teacher Follow-Up Survey*. National Center for Education Statistics.

Kember, D. (1997). A review and reconceptualization of the research into academics' conception of teaching. *Learning and Instruction, 7*(3), 255–275. https://doi.org/10.1016/S0959-4752(96)00028-X.

Kember, D., & Kwan, K. (2000). Lecturers' approaches to teaching and their relationship to conceptions of good teaching. *Instructional Science, 28*(5), 469–490. https://doi.org/10.1023/A:1026569608656

Kogan, D. M. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research, 62*(2), 129–169. https://doi.org/10.3102/00346543062002129

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology, 100*(1), 96–104. https://doi.org/10.1037/0022-0663.100.1.96

Kunnan, A. J. (2005). Language assessment from a wider context. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 779–794). Lawrence Erlbaum.

Lam, C. C., & Lipstone, J. (2007). Teachers' cultural differences: Case studies of geography. *Asia Pacific Education Review, 8*(2), 178-193. https://doi.org/10.1007/BF03026461

Lam, B. H., & Kember, D. (2006). The relationship between conceptions of teaching and approaches to teaching. *Teachers and Teaching: Theory and Practice, 12*(6), 693–713. https://doi.org/10.1080/13540600601029744

Leiter, M. P., & Maslach, C. (2005). A mediation model of job burnout. In A. S. G. Antoniou, & C. L. Cooper (Eds.), *Research companion to organizational health psychology* (pp. 544-564). Edward Elgar.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching, 40*, 243–259. https://doi.org/10.1017/S0261444807004363

Loh, C. H. E., & Liew, W. M. (2016). Voices from the ground: The emotional labour of English teachers' work. *Teaching and Teacher Education, 55*, 267-278. https://doi.org/10.1016/j.tate.2016.01.001

Lopez-Iñiguez, G., & Pozo, J. (2014). The influence of teachers' conceptions on their students' learning: Children's understanding of sheet music. *British Journal of Educational Psychology, 84*, 311–328. https://doi.org/10.1111/bjep.12033

Lunenburg, F. C. (2010). Creating a professional learning community. *National Forum of Educational Administration and Supervision Journal, 27*(4), 1-7.

Maloney, C., & Konza, D. (2011). A case study of teachers' professional learning: Becoming a community of professional learning or not? *Issues in Educational Research, 21*(1), 75–87.

Nias, J. (1996). Thinking about feeling: The emotions in teaching. *Cambridge Journal of Education, 26*(3), 293-306. https://doi.org/10.1080/0305764960260301

Oder, T., & Eisenschmidt, E. (2016). Teachers' perceptions of school climate as an indicator of their beliefs of effective teaching. *Cambridge Journal of Education, 46*. https://doi.org/10.1080/0305764X.2015.1045442

OECD. (2001). *Knowledge and skills for life: First results from PISA 2000*. Paris: OECD Publishing.

OECD. (2009). *Teaching and learning international survey (TALIS)*. Paris: OECD Publishing.

Perie, M., & Baker, D. P. (1997). *Job satisfaction among America's teachers: Effects of workplace conditions, background characteristics, and teacher compensation*. (ERIC Document Reproduction Service No. ED412 1810).

Pines, A., & Aronson, E. (1988). *Career burnout: Causes and cure*. Free Press.

Pratt, D. D., Arseneau, R., & Collins, J. B. (2001). Reconsidering "good teaching" across the continuum of medical education. *The Journal of Continuing Education in the Health Professions, 21*(2), 70-81. https://doi.org/10.1002/chp.1340210204

Qing-xue, L., & Jin-fang, S. H. (2007). An analysis of language teaching approaches and methods—Effectiveness and weakness. *US-China Education Review, 4*(1), 69-71.

Robinson, V. M. J., Lloyd, C., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership type. *Educational Administration Quarterly, 44*(5), 635–674. https://doi.org/10.1177/0013161X08321509

Rubie-Davies, C. M., Flint, A., & McDonald, L. G. (2012). Teacher beliefs, teacher characteristics, and school contextual factors: What are the relationships? *British Journal of Educational Psychology, 82*(2), 270–288. https://doi.org/10.1111/j.2044-8279.2011.02025.x.

Saavedra, A. R., & Opfer, V. D. (2012). Learning 21st-century skills requires 21st-century teaching. *Phi Delta Kappa, 94*(2), 8-13. https://doi.org/10.1177/003172171209400203

Sadeghi, K., & Khezrlou, S. (2016). The experience of burnout among English language teachers in Iran: Self and other determinants. *Teacher Development, 20*(5), 631-647. https://doi.org/10.1080/13664530.2016.1161660

Sadeghi, K., & Richards, J. (2015). The idea of English in Iran: An example from Urmia. *Journal of Multilingual and Multicultural Development, 37*(4), 419–434. https://doi.org/10.1080/01434632.2015.1093749

Sakurai, Y. (2012). Learners' perceptions of "good" foreign language teachers: A quantitative analysis between native and non-native teachers. *Electronic Journal of Foreign Language Teaching, 9*(1), 46–60.

Samuelowicz, K., & Bain, J. D. (2001). Revisiting academics' beliefs about teaching and learning. *Higher Education, 41*(3), 299–325. https://doi.org/10.1023/A:1004130031247

Serafini, F. (2002). Possibilities and challenges: The national board for professional teaching standards. *Journal of Teacher Education, 53*(4), 316–327. https://doi.org/10.1177/0022487102053004003

Solomon, D., Battistich, V., & Hom, A. (1996). Teacher beliefs and practices in schools serving communities that differ in socioeconomic level. Paper presented at the American Educational Research Association annual meeting, New York, NY.

Talbert, J. E., & Milbrey, W. M. (2002). Professional communities and the artisan model of teaching. *Teachers and Teaching: Theory and Practice, 8*, 325–343. https://doi.org/10.1080/135406002100000492

Tan, P. L. (2006). Approaches to learning and learning values: An investigation of adult learners in Malaysia (PhD dissertation). Queensland University of Technology. Retrieved from http://www.cybertesis.net/click?tu=http%3A%2F%2Feprints.qut.edu.au/16295/1/Po_Li_Tan_Thesis.pdf

Tavakoli, M., & Baniasad-Azad, S. (2016). Teachers' conceptions of effective teaching and their teaching practices: A mixed-method approach. *Teachers and Teaching: Theory and Practice, 23*(6), 674-688. https://doi.org/10.1080/13540602.2016.1218326

Taylor, D. L., & Tashakkori, A. (1995). Decision participation and school climate as predictors of job satisfaction and teachers' sense of efficacy. *Journal of Experimental Education, 63*(3), 217-231. https://doi.org/10.1080/00220973.1995.9943815

Thepa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357-385. https://doi.org/10.3102/0034654313483907

Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education, 61*, 47-59. https://doi.org/10.1016/j.tate.2016.10.001

Wenden, A. (2002). Learners' development in language learning. *Applied Linguistics, 23*(1), 32-55. https://doi.org/10.1093/applin/23.1.32

Winne, P. H., & Nesbit, J. C. (2010). The psychology of school performance. *Annual Review of Psychology, 61*, 653–678. https://doi.org/10.1146/annurev.psych.093008.100348

Woolfolk Hoy, A., Hoy, W. K., & Davis, H. A. (2009). Teachers' self-efficacy beliefs. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 627–653). Routledge.

The Dilemma of Metacognitive Intervention and EFL Listening: Is L1 a Panacea in EFL Contexts?

Ebrahim Fakhri Alamdari[1,*] & Leila Hosnbakhshan[1]

[1]Islamic Azad University, Qaemshahr Branch, Qaemshahr, Iran

Abstract
Listening has recently attracted the attention of both researchers and practitioners worldwide (Renandya & Hu, 2018), and research into L2 listening strategy use has recently tended to focus on metacognitive strategies (Lynch & Mendelsohn, 2020), This study investigated the comparative effect of L1/L2-mediated metacognitive intervention (MI) on the IELTS listening comprehension performance and metacognitive awareness of English as a foreign language (EFL) learners in Iran. The participants were 540 upper-intermediate EFL listeners in three groups, ranging from 17 to 28 years of age. The experimental groups (Ex1=180 / Ex2=180) went through a guided lesson plan in metacognition in English and Persian for twelve weeks, which focused on planning, monitoring, and evaluation. The control group (CG = 180), also instructed by the same teacher, listened to the same texts without any guided attention to the process. The MALQ and an actual IELTS test were used before and after the intervention to track the changes in metacognitive awareness and listening performance. The overall results showed that MI caused a considerable variance in the listening performance and the metacognitive awareness of learners in both experimental groups. The Post Hoc multiple comparison results of the three groups also illustrated that the medium for the delivery of the metacognitive intervention (L1) assisted the listeners in experimental group one, who went through L1-mediated metacognitive intervention, to outperform their peers in experimental group two, who were taught in L2, and the control group, who were taught conventionally.

1. INTRODUCTION

Listening comprehension, once a neglected skill (Renandya & Hu, 2018), is a complex process (Lynch & Mendelsohn, 2020). It facilitates the emergence of other language skills (Vandergrift & Goh, 2012) and is an essential skill for acquiring an additional language (Harding, Alderson, & Brunfaut, 2015; Wallace, 2020). Teaching of listening has received greater attention in recent years (Field 2008; Richards 2009), and the focus of teaching listening is on process rather than product (Graham & Santos, 2015; Santos & Graham, 2018).

[*]**Corresponding Author:** Ebrahim Fakhri Alamdari, Department of English, Islamic Azad University, Qaemshahr Branch, Qaemshahr, Iran. **Email**: fakhri.iau.ac@gmail.com
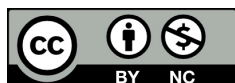
Metacognition, one of the most reliable predictors of learning (Vandergrift & Goh, 2012), has been widely recognized to have a crucial role in learning in general and L2/EFL listening, in particular. It is the ability of learners to control their thoughts and regulate their own learning, and can play an important role in learning to listen (Wenden, 1998). In addition, there is strong evidence that learners' metacognition can directly affect not only the process but also the outcome of their learning (Goh, 2018; Wenden, 1998). In the same vein, experts in the field of second language learning hold the view that learners' metacognitive awareness can contribute to their thinking and comprehension (Wenden, 1998), and can enhance a child's cognitive development, academic learning and language development, in general (Goh & Hu, 2014). One way to mitigate the cognitive demand of listening and facilitate the listening comprehension process for listeners is to use metacognitive intervention (Goh, 2008, 2018).

The use of L1 in ESL/EFL educational settings has also gained enormous recognition over the past two decades (Wach & Monroy, 2020), and many scholars have investigated the controversial impact of L1 on L2/EFL learning (Ellis & Shintani, 2014). Although the role of L1 in L2/EFL contexts has been revisited after decades of being underappreciated (De la Campa & Nassaji, 2009; Wach & Monroy 2020), there are still many scholars holding the view that the use of L1, as a pedagogic tool, could be interfering and debilitative and that L1 could undermine the L2/EFL learning process and may limit learners' L2 input and overshadow their exposure to L2 output (Swain & Lapkin, 2013; Tognini & Oliver, 2012). Other scholars, from an opposite angel, oppose the excessive use of L2, maintaining that it can lead to linguistic imperialism and may jeopardize the students' L1 as well as their culture. Thus, L1 must be deemed as a valuable pedagogic tool in L2/EFL educational settings, acting as a buffer against power relationships and potential cultural hazards in language classrooms (Stables & Wikeley, 1999). The use of L1 also gains support from a sociocultural perspective, claiming that it serves as an indispensable mediating tool assisting L2/EFL learners to complete cognitively demanding L2 tasks with more ease (Anton & Dicamilla,1999; Ellis & Shintani, 2014; Storch & Wigglesworth, 2003).

Although research on listening with regard to metacognitive interventions is broad in perspective, very few has ever focused on the potential impact of L1-mediated metacognitive intervention on L2 listening. In the light of this paucity in research, the present study focused on the comparative effect of L1/L2-mediated metacognitive intervention on the listening comprehension performance and metacognitive awareness of Iranian EFL learners.

2. REVIEW OF LITERATURE

Metacognitive Interventions and L2 Listening

Metacognition, widely thought to play a crucial role in a child's cognitive development, academic learning, and language development (Goh & Hu, 2014), empowers learners to control their thoughts, regulate their own learning and helps them learn how to listen (Goh, 2018; Wenden, 1998). Metacognitive intervention (MI), an overarching term primarily developed by Vandergrift and Goh (2012), encompasses both the strategy instruction and metacognitive instruction (Cross, 2015). It refers to pedagogical procedures empowering listeners to develop deeper metacognitive knowledge about themselves, the listening task, and appropriate strategies so as to enhance their awareness of the listening process (Vandergrift & Goh, 2012). MI also empowers listeners to plan, monitor, and evaluate their comprehension efforts as well as the progress of their overall listening development (Vandergrift & Goh, 2012).

In recent years, many scholars in the field of listening have developed models of metacognitive intervention (Vandergrift, 2004; Vandergrift & Goh, 2012; Vandergrift & Tafaghodtari, 2010). Many recent studies have applied these models in both ESL/EFL contexts to support the

importance of the MI for enhancing listening comprehension (Bozorgian, 2014; Bozorgian & Fakhri, 2018; Bozorgian & Muhammadpour, 2020; Bozorgian, Yaqubi, & Muhammadpour, 2020, Maftoon & Fakhri, 2020; Vandergrift & Tafaghodtari, 2010) as well as raising metacognitive awareness (Bozorgian & Fakhri, 2018; Maftoon & Fakhri, 2020; Mahdavi & Miri, 2017; Tanewong, 2018; Goh & Hu, 2014). The following studies are among the most recent ones reflecting the vital role of MI in enhancing listening comprehension and raising metacognitive awareness.

Bozorgian and Fakhri (2018) examined the effect of metacognitive intervention on advanced Iranian EFL learners' multimedia listening and their metacognitive awareness in listening. The data were collected through 180 Iranian advanced learners in three groups, two experimental and one control group. Multimedia listening tests and MALQ were to collect data for the study. The findings of the study demonstrated that metacognitive intervention enhanced learners' multimedia listening as well as their metacognitive awareness.

In a recent MI-based study, Maftoon and Fakhri (2020) also investigated the impact of metacognitive intervention on the listening performance and metacognitive awareness of 60 Iranian EFL learners. They used MALQ and a listening test to collect data for their study. The results were in favor of metacognitive intervention, enhancing listening performance and raising learners' metacognitive awareness.

In the most recent study, Bozorgian et al. (2020) investigated the effect of the metacognitive intervention on the listening performance and metacognitive awareness of 136 Iranian upper-intermediate EFL learners with low working memory capacity. IELTS listening tests and the MALQ were used before and after the intervention. The results depicted that the experimental group outperformed the control group in both IELTS listening and metacognitive awareness.

The Myth of L1 in L2/EFL Contexts

The ongoing debate over the use of L1 in L2/EFL classrooms has recently gained unprecedented recognition in educational settings worldwide (Shin et al., 2019). This controversy is well reflected in various academic publications, highlighting the role of L1, as a valuable pedagogic resource, in ESL/EFL classrooms (Butzkamm, 2003; Hall & Cook, 2012). This can also illustrate that the status of L1 has now been revisited after decades of being agonistically neglected by communicative approaches advocating L2- exclusivity inspired by natural approach (Krashen & Terrell, 1983) as well as arguments emanating from input, output and interaction (Long, 1996; Swain, 1985). As a compromise, Macaro et al. (2016) proposed three basic positions for the current status of L1 in L2 teaching: the first one is 'virtual position', denoting L2 exclusivity; the second one is 'maximal position', accepting occasional L1 use; and the third one is the 'optimal position', allowing the judicial use of the L1 as an important resource for L2 teaching.

Many scholars in the field of Applied Linguistics contend that L1 should be used in L2/EFL classrooms and further claim that they have convincing cognitive, sociolinguistic, sociocultural, psychological, and pedagogical reasons for their positions. From a cognitive point of view, they consider L1 users as cognitively sophisticated beings, who can use their L1 knowledge to perceive the concepts better in L2 (Butzkamm, 2003; Cook, 2001). From a socio-cognitive perspective, L1 is regarded as an invaluable tool to pool ideas, which can help mediate L2 learning and promote interaction in L2/EFL contexts (Anton & DiCamilla, 1999; Storch & Wigglesworth, 2003; Thoms et al., 2005). From a sociocultural perspective, L1 serves as a vital mediating tool assisting L2 learners to cope with and accomplish cognitively demanding L2 tasks more smoothly (Anton & Dicamilla,1999; Ellis & Shintani, 2014; Storch & Wigglesworth, 2003). From a psychological perspective, the use of L1 can reassure learners, boost their confidence, and ultimately help them

develop positive attitudes towards learning a second language (Littlewood, 2014). Pedagogically speaking, the use of L1 can assist teachers to establish rapport with learners more amicably and create a more friendly classroom atmosphere (Edstrom, 2006; Sali, 2014). In addition, L1 helps teachers save time, manage their classes more easily, keep pupils attended, and maintain classroom discipline (Auerbach, 2016).

Although there has been no severe backlash against the use of L1 in L2 settings in recent years (Cummins, 2007), and L1 is currently regarded as a scaffolding tool for the development of L2 (Auerbach, 2016), there are many who still cast doubt on the use of L1 in L2 instruction. They contend that the use of L1 may have deleterious impacts on L2, as it practically diminishes the quantity of comprehensible L2 input, ultimately hampering or interfering with L2 learning processes. In fact, they strive to promote the idea that L2 instruction should take place away from any L1- induced interference (Miles, 2004; Swain & Lapkin, 2013; Tognini & Oliver, 2012).

Overall, SLA researchers have been at pains to prove that L1 is not just a setback to L2 learning but a resource for learners to facilitate their use and learning of an L2 (Ellis & Shintani, 2014). Notwithstanding all the benefits and pitfalls, the following two studies are among the few relevant studies investigating the effect of L1 on listening comprehension performance of listeners in both EFL/ESL contexts.

In a research study investigating the impact of L1 on the use of two particular listening strategies, syntactic cues and prosodic cues, Harley (2000) concluded that Chinese and Polish EFL learners with various levels of proficiency tended to seek prosodic assistance from L1. In other words, they relied on syntax to reconstruct prosodic cues once they encounter confusing sentences. In another study investigating the impact of L1-mediated instruction of listening strategies and its effect on listening comprehension in L2, Bozorgian and Pillay (2013) taught five listening strategies in L1 to sixty Iranian EFL learners over 14 weeks. The results of the study were in favor of L1 instruction, leading to improvements in EFL listening.

It is worth noting that none of the studies mentioned above has ever focused on the comparative effect of L1/L2-mediated metacognitive intervention on the listening comprehension performance and metacognitive awareness of EFL learners. Thus, in an attempt to reach more tangible and consistent findings regarding the efficacy of metacognitive interventions, the present study strove to investigate the matter through the following research questions:

1. Does L1-mediated metacognitive intervention have any effect on the IELTS listening comprehension performance of Iranian EFL listeners?
2. Does L1-mediated metacognitive intervention have any effect on the metacognitive awareness of Iranian EFL listeners?

3. METHOD

Participants

The participants for this study were chosen from among all the available EFL learners preparing for IELTS at an English language institute in Iran. Having screened the EFL listeners through Oxford Placement Tests (OPTs) over the period of three consecutive terms, the researchers chose 540 upper- intermediate male and female participants, who were between 17 and 24 years of age, and randomly assigned them to two experimental (Ex1 = 180 / Ex2 = 180) and a control (CG = 180) group prior to implementing the intervention programs. Consent forms were obtained from all participants, and the participants were clarified with regard to the nature and purpose of the study. The researcher carrying out the intervention programs for both experimental groups in this study was a Ph.D. holder in TEFL, who was teaching in the same language institute for more than 20 years and was quite familiar with both educational contents and MI.

Instruments

Three instruments were used to collect data for the research questions of this study.

Oxford Placement Test (Allan, 2004) was used to screen the participants in terms of homogeneity at the outset of the study. It also served as a criterion to estimate the reliability as well as the concurrent validity of the IELTS listening tests in this study. The test had high Cronbach's alpha consistency reliability, .94 (Larson-Hall, 2010).

Metacognitive Awareness Listening Questionnaire (Vandergrift et al., 2006) was used to measure EFL learners' metacognitive awareness and their perceived use of metacognitive strategies at the beginning and end of the study. The 6-point Likert-scale MALQ comprises 21 items covering five factors: problem-solving, planning and evaluation, mental translation, person knowledge, and directed attention. Its internal reliability estimates range from .68 to .78 (Vandergrift et al., 2006) and it also enjoys high validity (Goh & Hu, 2014).

Cambridge IELTS 14, published by Cambridge University Press and UCLES (2019), was used to assess the learners' listening performance before and after the intervention. The Cronbach's alpha reliability and the concurrent validity of the IELTS listening test were both high, 0.91 and 0.86, respectively (Larson-Hall, 2010).

Procedures

Having received the approval from the headquarter of the English language institute and screened the participants through OPT, the researchers administered the IELTS listening test and the MALQ to all participants as pre- test at the outset and post-test at the end of the intervention to compare and assess their initial and final performances on both and further investigate the probable effect(s) of MI. The metacognitive intervention was presented to the participants in three phases:

In phase one, weeks one to four, the MI dealt with planning, focusing on advanced organizers, directed and selective attention, and learning management. All these concepts were fully clarified to the participants through ample examples and explanations. The researcher then helped the learners concentrate on what they were listening to and tried to establish information linkage during listening. As regards directed attention, the researcher told the listeners to overlook irrelevant distracters and keep their attention focused on what was happening in the listening. In the case of selective attention, the researcher wanted the listeners to focus on the given topic and identify the key words prior to the listening. For learning management, the researcher advised the listeners to understand various contexts while listening and strive to adapt themselves to any new circumstances they encountered throughout the listening. In this phase, the researcher advised the learners to frame their mind to understand the audio text. The listeners were advised to keep their attention focused on what the speakers were talking about. In the second phase, weeks five to seven, the MI concerned with monitoring, concentrating on comprehension, auditory, and double-checking monitoring. The researcher elaborated on the definition of monitoring and focused on comprehension monitoring strategies, in particular. The researcher advised the learners to translate any odd words to see if they sounded right and try to put everything together, as understanding one thing could lead to the understanding of another. Through auditory monitoring, the listeners learned how sounds made sense and through double-checking. They realized how to check comprehension throughout listening. In the final phase, weeks eight to ten, the MI covered evaluation, concentrating on performance evaluation, strategy evaluation, and problem identification. Having defined the concept of evaluation and gone through the listening activity, the researcher emphasized the power of evaluation, and reiterated that performance evaluation in listening input could fill out where they lack understanding in listening. With regard to strategy

evaluation, the researcher further emphasized the use of learning strategy before and while listening, and reiterated that strategy evaluation in listening input could develop the use of listening strategy when the need is felt. Having familiarized the learners with problem identification strategy, the researcher recommended the listeners to pinpoint the areas where misunderstanding happened during listening check and recheck. Then the listeners instructed how to overcome their previous listening barriers.

The Intervention Programs

The intervention for the experimental groups was based on Vandergrift and Goh's "Pedagogical Cycle" (2012). The pedagogical cycle comprised five stages including planning/predicting, first verification, second verification, final verification, and reflection (Vandergrift & Goh, 2012) and was utilized as a process-based approach to raise the listeners' metacognitive awareness in listening. The participants in the experimental groups one and two participated in a twelve-week metacognitive instruction program, twice a week, each about 90 minutes. Each session, the participants listened to a different oral text, which was aligned with the content of the intervention program. It should be noted that the medium for the delivery of MI to learners in experimental groups one and two were L1 (Persian) and L2 (English), respectively. Having implemented the intervention, the researcher administered the posttest and the MALQ to explore the probable effect(s) of the intervention.

The participants in the control group were exposed to the traditional approach, comprising pre-, while-, and post-listening stages for every listening task. They listened to the same texts but were not engaged in any formal prediction activity, nor were they given the chance to experience the process of listening through discussing, predicting, or monitoring their comprehension with their peers. In pre-listening stage, they were given a warm-up related to the topic of the listening task in order that they might activate their prior knowledge for improved listening comprehension. In the while-listening stage, the learners listened to the CD to complete a task that was aligned with the pedagogical contents presented to them in each unit. They were allowed to listen to the aural text as many times as the instruction in each unit required them. Having listened to the aural text, the learners read their answers one by one to get them checked. To help the learners solve their potential problems, the teacher played the CD once more to deal with any ambiguities the listeners might have encountered throughout the listening task. In the post-listening stage, the teacher asked one or two learners to give a summary and engaged the class in a discussion in order to confirm their comprehension of the text. Also, there was no discussion of strategy use, nor were the students engaged in any formal reflection on their approach to listening.

Data Analysis

The Statistical Package for the Social Sciences (SPSS) version 22 was employed to analyze the IELTS listening test and MALQ scores to respond to the research questions. First, the equality of variance and the distribution of the data were analyzed. Then, due to the non- normality of the data set, the data collected from the two groups were analyzed using the Kruskal-Wallis test for pre- and post-tests.

Table 1: Descriptive statistics of the overall scores of the three groups for the pre- and post-tests of IELTS listening

Tests	Groups	Minimum	Maximum	Mean	Std. Deviation	N
Pre-tests	Ex1	7.00	18.00	12.53	2.54	180
	Ex2	10.00	15.00	12.52	1.26	180
	CG	7.00	18.00	12.62	2.58	180
Post-tests	Ex1	9.00	20.00	16.15	2.16	180
	Ex2	11.00	20.00	14.56	1.55	180
	CG	8.00	18.00	12.57	2.43	180

Table 2: Kruskal Wallis test results comparing the three groups in the pre- and post-tests of IELTS listening

Scores	Groups	N	Mean Rank	Chi-square	df	Sig.
Pre-tests	Ex1	180	262.41			
	Ex2	180	275.97	.776	2	.679
	CG	180	273.12			
Post-tests	Ex1	180	375.39			
	Ex2	180	267.70	161.24	2	.000
	CG	180	168.41			
	The effect size		.30			

4. RESULTS

The reliability indices of Cronbach's alpha reported in the pre- and post-tests for experimental (EG1: $\alpha = .89$, $\alpha = .87$; EG2: $\alpha = .85$, $\alpha = .88$) and control ($\alpha = .87$, $\alpha = .84$) groups of IELTS listening and MALQ, experimental (EG1: $\alpha = .90$, $\alpha = .89$; EG2: $\alpha = .88$, $\alpha = .89$) and control ($\alpha = .85$, $\alpha = .87$), were moderate (Larson-Hall, 2010). The pretests of listening and MALQ indicated the listeners' base-line information on language competence level and metacognitive awareness.

Research Question 1

The first research question focuses on whether the L1-mediated metacognitive intervention has any significant effect on EFL learners' listening performance. The answer to this question is affirmative.

The results of the descriptive statistics (Table 1), Ex1 (M = 16.15; SD = 2.16), Ex2 (M = 14.56; SD = 1.55), and the CG (M = 12.75; SD = 2.43), and Kruskal-Wallis test (Table 2) for the post-test of IELTS listening indicated that the participants in Ex1 and Ex2 outperformed their peers in the CG. In fact, the mean rank score obtained by Ex1 (375.39) exceeds the mean score obtained by Ex2 (267.70), which is, in turn, higher than the mean score of CG (168.41). The Chi-square value (161.24) is more than the critical value for the df of 2 and the p-value of .05, suggesting that the differences in mean scores among the three groups are statistically significant after the intervention. The effect size of .30 is moderate (Cohen, 1988). The p-. value of .00 (.00 < .05) also implies that the intervention programs did lead to a great variance in the listening performance of learners in both experimental groups, suggesting that the medium of instruction can be effective, and can lead to variance in learners' listening performance.

Table 3: Descriptive statistics of the overall scores of the three groups for the pre- and post-tests of MALQ

Tests	Groups	Minimum	Maximum	Mean	Std. Deviation	N
	Ex1	65.00	97.00	82.16	7.73	180
Pre-tests	Ex2	61.00	110.00	81.85	9.42	180
	CG	64.00	99.00	80.87	7.59	180
	Ex1	80.00	120.00	94.81	7.53	180
Post-tests	Ex2	61.00	116.00	91.26	10.17	180
	CG	52.00	100.00	79.95	9.05	180

Table 4: Kruskal Wallis test results comparing the three groups in the pre- and post-tests of MALQ

Scores	Groups	N	Mean Rank	Chi-square	df	Sig.
	Ex1	180	289.60			
Pre-tests	Ex2	180	264.03	4.196	2	.123
	CG	180	257.87			
	Ex1	180	361.41			
Post-tests	Ex2	180	305.68	188.04	2	.000
	CG	180	144.41			
The effect size			.33			

Research Question 2

The second research question asks whether the L1-mediated metacognitive intervention has any effect on the metacognitive awareness of Iranian EFL listeners. The answer to this question is also affirmative

Both the descriptive statistics (Table 3), Ex1 (M = 94.81; SD = 7.53), Ex2 (M = 91.26; SD = 10.17), and the CG (M = 79.95; SD = 9.05) and Kruskal-Wallis test (Table 4) demonstrated that the participants in both experimental groups outperformed their peers in the control group in the post-test of strategy questionnaire. The Chi-square value (188.04) is more than the critical value for the df = 2 and the p-value = .05, suggesting that the differences in mean scores among the three groups are statistically significant after the intervention. The effect size of .33 is moderate (Cohen, 1988). The p- value = .00 (.00 < .05) also implies that the intervention programs, the medium of instruction, did lead to a significant variance in the metacognitive awareness of learners in Both experimental groups after the intervention.

To find out the exact differences among the three groups in terms of IELTS listening performance and metacognitive awareness, the researchers had to utilize a post-hoc Tukey's HSD Test (Table 5) in the light of the fact that this kind of test can pinpoint where the exact differences among the three groups lie.

Table 5: Post Hoc Tukey's HSD test for the post-tests of listening and MALQ

Tests	Groups	N	Mean Difference	Std. Error	Sig.
	Ex1 vs. Ex2	180	1.58	.219	.000
Listening	Ex1 vs. CG	180	3.40	.219	.000
	Ex2 vs. CG	180	1.81	.219	.000
	Ex1 vs. Ex2	180	3.54	.947	.000
MALQ	Ex1 vs. CG	180	13.85	.947	.000
	Ex2 vs. CG	180	11.31	.947	.000

As regards IELTS listening, the results show that there was a statistically significant difference between Ex1 and Ex2 in terms of their listening performance after the intervention programs. This can further suggest that metacognitive instruction through L1 proved to be more effective and led to a greater variance in the listening performance of learners in Ex1. In pair two, the results also show that there was a statistically significant difference between the two groups, Ex1 and CG, in terms of their listening performance after the intervention, as the p-value is less than .05 (.006 < .05). This implies that the intervention program designed for the participants in Ex1 resulted in a variance in their listening performance, compared with their peers in the CG, who were taught conventionally without any attention to the process. The comparison, in pair three, lies between Ex2 and CG. The results revealed that there was a statistically significant difference between the two groups in terms of their listening performance after the intervention, as the p-value is less than .05 (.00 < .05). This result also suggests that metacognitive instruction through L2 also helped the learners in Ex2 improve their listening performance, and led to a great variance in the overall post-test results, compared with the participants in the CG, for whom the conventional teaching of listening was in practice.

Turning to metacognitive awareness, the results indicate that there was a statistically significant difference between Ex1 and Ex2 in terms of their metacognitive awareness after the intervention program, suggesting that metacognitive instruction through L1 proved to be more effective and led to a greater variance in the metacognitive awareness of learners in Ex1. In pair two, there was a statistically significant difference between Ex1 and CG in terms of their metacognitive awareness after the intervention, as the p-value is less than .05 (.006 < .05). This implies that the intervention program designed for the participants in Ex1 resulted in a variance in their metacognitive awareness, compared with their peers in the CG, who were taught conventionally without any attention to the process. The comparison between Ex2 and CG in pair three revealed that there was a statistically significant difference between the two groups in terms of their metacognitive awareness after the intervention, as the p-value is less than .05 (.000 < .05). This result also suggests that metacognitive instruction through L2 also helped the learners in Ex2 raise their metacognitive awareness, and led to a great variance in the overall post-test results, compared with the participants in the CG, for whom the conventional teaching of listening was in practice.

5. DISCUSSION

Using L1 plays a clearly complex role in L2 learning, and there is conspicuous paucity in research investigating the facilitative or debilitative effect of L1 on actual learning (Ellis & Shintani, 2014). The focus of research into L2 listening strategy use has recently shifted to metacognitive strategies (Lynch & Mendelsohn, 2020). This cohort study was to examine the comparative effect of L1/L2-mediated metacognitive intervention on IELTS listening

comprehension performance and metacognitive awareness of Iranian EFL learners. The overall result of the study is in favor of L1 use. In other words, the medium of delivering the metacognitive intervention (L1) assisted the listeners in experimental group one to outperform their peers in experimental group two and the control group in both the listening performance and the metacognitive awareness.

As regards listening performance, the findings of this study are consistent with those of (Bozorgian, 2014; Bozorgian & Fakhri, 2018; Bozorgian & Muhammadpour, 2020; Bozorgian, Yaqubi, & Muhammadpour, 2020, Maftoon & Fakhri, 2020; Vandergrift & Tafaghodtari, 2010), underscoring the indispensable role metacognitive intervention can play in the enhancement of listening skill. This further suggests that should listening be taught in tandem with metacognitive interventions through a process-based approach, better outcomes could be expected (Maftoon & Fakhri, 2020). The results of this study are, nevertheless, at odds with those of two other studies (Chen & Huang, 2011; Rahimi & Katal, 2013) investigating the potential benefits of MI and MALQ, leading to no immediate outcome in terms of improvements in listening. The length of MI, contextual factors, and the listeners' level of communicative competence could be the potential causes of this mismatch (Bozorgian, 2014; Maftoon & Fakhri, 2020),

With regard to metacognitive awareness, the results are congruous with those of (Bozorgian & Fakhri, 2018; Bozorgian & Muhammadpour, 2020; Maftoon & Fakhri, 2020; Mahdavi & Miri, 2017; Tanewong, 2018; Goh & Hu, 2013), all confirming the contribution of process-based MI to metacognitive awareness. This result, however, empirically contradicts Bozorgian (2014) who reported no conspicuous improvements in learners' metacognitive awareness after the implementation of MI. Listeners' inadequate knowledge about metacognitive strategies, their failure to perceive and apply these strategies in practice, and their lack of understanding the function of these factors are the main factors to consider in the case of this discrepancy (Bozorgian, 2014; Maftoon & Fakhri, 2020). Perhaps, this puzzle, the discrepancies in results and the potential factors behind them, highlights the urgent need for the use of L1 in both ESL/EFL contexts.

Using L1, a cognitive resource for L2 learners, is one of the primary means through which learners can mediate L2 learning (Ellis & Shintani, 2014). The tendency for L1 use has grown in strength in recent years, and it is obvious that the pendulum has swung in favor of L1 in applied linguistics (Hall & Cook, 2012). Turning to L1, the result of this study corroborates those of Harley (2000) and Bozorgian and Pillay (2013), who used L1 as the medium of instruction in EFL context. These findings also suggest that the use of EFL learners' L1, as the medium of instruction, can empower learners to better perceive complex listening strategies, which might look perplexing once presented in a foreign language, especially for listeners with limited linguistic competency. This is in line with Graham and Macaro (2008), maintaining that the intervention can bring about enhanced listening proficiency. In addition, this notion can be substantiated by Field (2008), maintaining that the use of L1 can help EFL learners compensate for their limited linguistic competence, inadequate vocabulary repertoire, limited listening exposure, and working memory capacity.

Theoretically, Macaro (2009), endorsing the learning contribution of L1 to L2, proposes a framework demonstrating that the use of L1 can result in enhanced learning in L2 through three distinctive sources. To begin with, he asserts that, psycho-linguistically, predicting, processing and storing knowledge are all tightly merged with the cognitive theory used in both L1 and L2 language learning through interaction in both short and long-term memory. Secondly, he argues that socio-cultural theory backs L1 assistance in L2 learning and stresses the notion that both think aloud and engaging in mental commentaries taking place in L1 can contribute to L2 learning. He

ultimately justifies the issue through the lens of code-switching theory and argues that L1 can facilitate the process of L2 learning through linguistic styles (both formal and informal) in real life interactions.

Commenting on the extent to which L1 should be incorporated into teacher talk. Macaro (2009) strongly argues in favor of intra- sentential rather than inter-sentential codeswitching and strongly advocates the notion that teachers ought not to produce the whole sentences in L1, but should use it strategically for the clarification of essential words and lexical strings.

The overall results of the present study also revealed a significant difference between Ex1, Ex2, and CG in terms of their metacognitive awareness, which was achieved through the intervention programs in this study. However, a closer look at the results of the two experimental groups (Ex1 & Ex2) in this study indicated that the medium of metacognitive instruction did lead to a superiority of one group over the other. In other words, there was a significant difference between the two experimental groups (Ex1 & Ex2) in terms of their listening performance and metacognitive awareness after the implementation of the intervention programs, suggesting that the listeners in Ex1 who went through L1-mediated metacognitive intervention outperformed their peers in Ex2, who went through a metacognitive intervention program in L2. This result might not have, otherwise, been achieved. This can further substantiate Mendelsohn (1995), pointing out that first language strategies ought to be unlocked so that learners can implement them automatically in L2 learning.

With regard to the significance of using L1 in L2 learning development, research findings (See Carless, 2008) illustrate that the belief system of language teachers considering L1 as interference (See Kellerman, 1995) in L2 learning development has been transformed in the last two decades (Ellis & Shintani, 2014). Consequently, experienced teachers, inexperienced teachers, and teacher trainers view L1 not only as a constructive means to scaffold learning but also as an effective means of classroom management (Littlewood & Yu, 2009). This can further support the inclusion of L1 into curriculum as well as classroom syllabi.

6. CONCLUSION

This study delved into the comparative effect of L1/L2-mediated metacognitive intervention on IELTS listening comprehension performance and metacognitive awareness of Iranian upper-intermediate EFL learners. Returning to the focal inquiry in this study as to whether L1 is a panacea in EFL contexts, the overall result of the study is amazingly in favor of L1. Given the obtained result, it can be concluded that L1 can be used as an aid in EFL classes assisting teachers in establishing rapport and solidarity, maintaining and facilitating communication throughout a lesson, conveying meaning, mitigating the learners' anxiety, and clarifying the vague points throughout teaching (Auerbach, 2016; Edstrom, 2006; Sali, 2014). Students should bear in mind that L1 ought to be used judiciously for scaffolding and peer learning, contributing to shape their L2/EFL knowledge. Another word of caution is that L1 must not overshadow L2 in EFL contexts, which can cause interference and laziness. The major contribution of this study lies in its initiative to have compared the effect of two various media of metacognitive intervention, Persian and English, on the IELTS listening performance and metacognitive awareness of EFL listeners, as none in the literature has ever sought to investigate the impact of these two media on the listening performance and metacognitive awareness of ESL / EFL learners. With regard to pedagogical implications, the findings underscore the urge for heightening the learners' metacognitive awareness and use, especially in EFL settings. One way to promote their awareness is to incorporate awareness– raising tasks in EFL instructional materials, whether in L1 or in L2, to draw their attention to different learning strategies making them more motivated and ultimately

self-regulated throughout language learning process. The findings, therefore, shed light on the notion that devoting enough time to listening activities aligned with an L1/L2-mediated MI, can alter listeners' outlook towards learning in general and listening input in particular. Thus, it is incumbent upon curriculum developers and textbook writers to incorporate and highlight listening strategies in textbooks and consider devoting ample time to each listening task so as to promote the significance of teaching listening strategies in the classroom. Regarding future research directions, while MI successfully improved EFL learners' listening performance, it is not known to what extent and in what ways metacognitive instruction contributed to the listening improvement. Future research designs necessitate considering each metacognitive strategy separately so that the relative contribution of each to listening can be clearly pinpointed. Additionally, given the complex and largely internal nature of metacognitive strategy use, future research needs to investigate metacognitive awareness as measured by the MALQ and actual metacognitive strategy use through the use of stimulated recall or think- aloud protocols. There is an urgent need to interview students to find out which medium of instruction, L1 or L2, sounds more effective to them. Last but not least, this study focused on learners and tried to investigate their listening performance and metacognitive awareness through a process-based approach. Future research can, likewise, investigate the teachers' views on teaching listening and the issues they face during teaching this basic skill.

Disclosure statement
No potential conflict of interest to be reported.

Funding
This work was self-funded.

Data availability statement
Given the privacy of research participants, the data supporting the findings of this study are available on request from the corresponding author.

References

Allan, D. (2004). *Oxford placement tests 2: Test pack*. Oxford University Press.

Antón, M., & DiCamilla, F. J. (1999). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *The Modern Language Journal, 83*(2), 233–247. https://doi.org/10.1111/0026-7902.00018

Auerbach, E. (2016). Reflections on Auerbach (1993), Reexamining English only in the ESL classroom. *TESOL Quarterly, 50*, 936–939. https://doi.org/10.1002/tesq.310

Bozorgian, H. (2014). The role of metacognition in the development of EFL learners' listening skill. *International Journal of Listening, 28*(3), 149–161. https://doi.org/10.1080/10904018.2013.861303.

Bozorgian, H., & Fakhri Alamdari, E. (2018). Multimedia listening comprehension: Metacognitive instruction or metacognitive instruction through dialogic interaction. *ReCALL, 30*(1), 131–152. https://doi.org/10.1017/S0958344016000240

Bozorgian, H., & Muhammadpour, M. (2020). Metacognitive intervention: High WMC learners' listening performance and metacognitive awareness. *Foreign Language Research Journal, 9*(4), 1055–1084. https://doi.org/10.22059/JFLR.2019.273053.592.

Bozorgian, H., & Pillay, H. (2013). Enhancing foreign language learning through listening strategies delivered in L1: An experimental study. *International Journal of Instruction, 6*(1), 105–122.

Bozorgian, H., Yaqubi, B., & Muhammadpour, M. (2020). Metacognitive intervention and awareness: Listeners with low working memory capacity. *International Journal of Listening*. https://doi.org/10.1080/10904018.2020.1857764

Butzkamm, W. (2003). We only learn language once. The role of the mother tongue in FL classrooms: Death of a dogma. *The Language Learning Journal, 28*(1), 29–39. https://doi.org/10.1080/09571730385200181

Carless, D. (2008). Student use of the mother tongue in the task-based classroom. *ELT Journal, 62*(4), 331–338. https://doi.org/10.1093/elt/ccm090

Chen, C. C., & Huang, C. T. (2011). Raising EFL college students' metacognitive awareness about listening. Paper presented at the NYS TESOL 41st Annual Conference, Melville, NY.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review, 57*(3), 402–423. https://doi.org/10.3138/cmlr.57.3.402

Cross, J. (2015). Metacognition in L2 listening: Clarifying instructional theory and practice. *TESOL Quarterly, 49*(4), 883–892. https://doi.org/10.1002/tesq.258

Cummins, J. (2007). Rethinking monolingual instructional strategies in multilingual classrooms. *Canadian Journal of Applied Linguistics, 10*(2), 221–240. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=20680862

De la Campa, J. C., & Nassaji, H. (2009). The amount, purpose, and reasons for using L1 in L2 classrooms. *Foreign Language Annals, 42*(4), 742–759. https://doi.org/10.1111/j.1944-9720.2009.01052.x.

Edstrom, A. (2006). L1 use in the L2 classroom: One teacher's self-evaluation. *Canadian Modern Language Review, 63*(2), 275–292. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=18369910

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London: Routledge.

Field, J. (2008). Guest editor's introduction: Emergent and divergent: A view of second language listening research. *System, 36*, 2–9. https://doi.org/10.1016/j.system.2008.01.001

Goh, C. C. (2008). Metacognitive instruction for second language listening development: Theory, practice, and research implications. *RELC Journal, 39*(2), 188–213. https://doi.org/10.1177/0033688208092184

Goh, C. C. (2018). Metacognition in second language listening. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & Sons. https://doi.org/10.1002/9781118784235.eelt0572

Goh, C. C., & Hu, G. (2014). Exploring the relationship between metacognitive awareness and listening performance with questionnaire data. *Language Awareness, 23*(3), 255–269. https://doi.org/10.1080/09658416.2013.769558

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning, 58*(4), 747–783. https://doi.org/10.1111/j.1467-9922.2008.00478.x

Graham, S., & Santos, D. (2015). *Strategies for second language listening: Current scenarios and improved pedagogy*. Palgrave Macmillan.

Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning: State of the art. *Language Teaching, 45*(3), 271–308. https://doi.org/10.1017/S0261444812000067

Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing, 32*(3), 317–336. https://doi.org/10.1177/0265532214564505

Harley, B. (2000). Listening strategies in ESL: Do age and L1 make a difference? *TESOL Quarterly, 34*(4), 769–777. https://doi.org/10.2307/3587790

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual Review of Applied Linguistics, 15*, 125–150. https://doi.org/10.1017/S0267190500002658

Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Pergamon Press.

Littlewood, W. (2014). Communication-oriented teaching: Where are we now? Where do we go from here? *Language Teaching, 47*(3), 249–362. https://doi.org/10.1017/S0261444812000134.

Littlewood, W., & Yu, B. H. (2011). First language and target language in the foreign language classroom. *Language Teaching, 44*, 64–77. https://doi.org/10.1017/S0261444809990310

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Lynch, T., & Mendelsohn, D. (2020). Listening. In N. Schmitt & M. P. H. Rodgers (Eds.), *An introduction to applied linguistics* (pp. 223–239). Routledge.

Macaro, E. (2009). Teacher use of codeswitching in the second language classroom: Exploring 'optimal' use. In M. Turnbull & J. Dailey-O'Cain (Eds.), *First language use in second and foreign language learning* (pp. 35–49). Multilingual Matters.

Macaro, E., Graham, S., & Woore, R. (2016). *Improving foreign language teaching: Towards a research-based curriculum and pedagogy*. Routledge.

Maftoon, P., & Fakhri Alamdari, E. (2020). Exploring the effect of metacognitive strategy instruction on metacognitive awareness and listening performance through a process-based approach. *International Journal of Listening, 34*(1), 1–20. https://doi.org/10.1080/10904018.2016.1250632

Mahdavi, N., & Miri, M. (2017). Co-shaping metacognitive awareness and developing listening comprehension through process-based instruction. *International Journal of Listening, 33*(1), 53–70. https://doi.org/10.1080/10904018.2016.1260454.

Mendelsohn, D. (1995). Applying learning strategies in the second/foreign language listening comprehension lesson. In D. Mendelsohn & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 132–150). Dominie Press.

Rahimi, M., & Katal, M. (2013). The impact of metacognitive instruction on EFL learners' listening comprehension and oral language proficiency. *Journal of Teaching Language Skills, 5*(2), 69–90. https://doi.org/10.22099/jtls.2013.1555

Renandya, W. A., & Hu, G. (2018). L2 listening in China: An examination of current practice. In A. Burns & J. Siegel (Eds.), *International perspectives on teaching the four skills in ELT* (pp. 37–50). Palgrave Macmillan.

Richards, J. C. (2009). *Teaching listening and speaking*. SEAMEO Regional Language Centre.

Sali, P. (2014). An analysis of the teachers' use of L1 in Turkish EFL classrooms. *System, 42*, 308–318. https://doi.org/10.1016/j.system.2013.12.021

Santos, D., & Graham, S. (2018). What teachers say about listening and its pedagogy: A comparison between two countries. In A. Burns & J. Siegel (Eds.), *International perspectives on teaching the four skills in ELT* (pp. 21–35). Palgrave Macmillan.

Shin, J., Dixon, Q. L., & Choi, Y. (2019). An updated review on use of L1 in foreign language classrooms. *Journal of Multilingual and Multicultural Development, 41*(5), 406–419. https://doi.org/10.1080/01434632.2019.1684928

Stables, A., & Wikeley, F. (1999). From bad to worse? Pupils' attitudes to modern foreign languages at ages 14 and 15. *The Language Learning Journal, 20*(1), 27–31. https://doi.org/10.1080/09571739985200231

Storch, N., & Wigglesworth, G. (2003). Is there a role for the use of the L1 in an L2 setting? *TESOL Quarterly, 37*(4), 760–770. https://doi.org/10.2307/3588224

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

Swain, M., & Lapkin, S. (2013). A Vygotskian sociocultural perspective on immersion education: The L1/L2 debate. *Journal of Immersion and Content-based Language Education, 1*(1), 101–129. https://doi.org/10.1075/jicb.1.1.05swa.

Tanewong, S. (2018). Metacognitive pedagogical sequence for less-proficient Thai EFL listeners: A comparative investigation. *RELC Journal, 50*(1), 86–103. https://doi.org/10.1177/0033688218754942

Thoms, J., Liao, J., & Szustak, A. (2005). The use of L1 in an L2 on-line chat activity. *Canadian Modern Language Review, 62*(1), 161–182. https://www.learntechlib.org/p/98046/

Tognini, R., & Oliver, R. (2012). L1 use in primary and secondary foreign language classrooms and its contribution to learning. In E. Alcón Soler & M. P. Safont Jordá (Eds.), *Discourse and learning across L2 instructional contexts* (pp. 53–78). Rodopi.

UCLES, C. U. P. A. (2019). *Cambridge IELTS 14 academic student's book with answers: Authentic examination papers*. Cambridge University Press.

Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics, 24*, 3–25. https://doi.org/10.1017/S0267190504000017

Vandergrift, L., & Goh, C. C. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Vandergrift, L., Goh, C. C. M., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language Learning, 56*(3), 431–462. https://doi.org/10.1111/j.1467-9922.2006.00373.x.

Vandergrift, L., & Tafaghodtari, H. M. (2010). Teaching L2 students how to listen does make a difference: An empirical study. *Language Learning, 60*(2), 470–497. https://doi.org/10.1111/j.1467-9922.2009.00559.x

Wach, A., & Monroy, F. (2020). Beliefs about L1 use in teaching English: A comparative study of Polish and Spanish teacher-trainees. *Language Teaching Research, 24*(6), 855–873. https://doi.org/10.1177/1362168819830422

Wallace, M. P. (2020). Individual differences in second language listening: Examining the role of knowledge, metacognitive awareness, memory, and attention. *Language Learning*. https://doi.org/10.1111/lang.12424.

Wenden, A. (1998). Metacognitive knowledge and language learning. *Applied Linguistics, 19*(4), 515–537. https://doi.org/10.1093/applin/19.4.515

Iranian Advanced EFL Learners' Perceptions of the Gravity of Their Peer Written Lexical Errors: The Case of Intelligibility and Acceptability

Musa Nushi[1][*] iD, Roya Jafari[2], and Masoumeh Tayyebi[3]

[1] Shahid Beheshti University, Tehran, Iran
[2] PhD Candidate, Tehran University, Tehran, Iran
[3] PhD Candidate, Islamic Azad university, Savadkooh Branch, Savadkooh, Iran

Abstract
Errors are an important feature of development in a second language as they indicate the state of learners' knowledge. This study investigates the gravity of the lexical errors made by Iranian advanced EFL learners from the perspective of their peer advanced EFL learners. Sixty advanced Iranian undergraduate students who were majoring in English Language and Literature at Shahid Beheshti University of Tehran, Iran, took part in this research. The participants, who were selected through purposive sampling, were given a questionnaire containing eleven lexically erroneous sentences extracted from their fellow advanced students' writings. They were required to judge those sentences in terms of their acceptability and intelligibility. The results indicated that the students considered mis-ordering as least acceptable type of errors (mean:1.78) and calque (i.e., translation from L1) the least intelligible type (mean: 2.46). The results further showed that there was a strong positive correlation between acceptability and intelligibility ratings of the errors by the advanced EFL learners, meaning that the more acceptable the errors, the more intelligible they were. The findings of this study can help improve our understanding of EFL learners' problems; they can also inform EFL teachers' instructional planning and remedial practices, especially in the English as an international language paradigm.

1. INTRODUCTION

Second language (L2) teachers face persistent challenges when dealing with students' linguistic errors (Hyland & Anan, 2006). However, errors are inevitable in the process of L2 learning and how to best address them has always been high on L2 teachers' agenda. For a long time, students' errors were not welcomed (Larsen-Freeman & Anderson, 2013) and they were to be avoided at all cost and immediately corrected by teachers if they occurred. With the advent of communicative language teaching approach; nevertheless, these negative perceptions towards errors began to change as scholars found that errors are unavoidably necessary to students' learning (Yang et al.,

[*] **Corresponding Author:** Musa Nushi, Department of English Language and Literature, Shahid Beheshti University, Tehran, Iran, **Email**: m_nushi@sbu.ac.ir

2016). This perspective brought about a shift in views towards errors and the way they should be dealt with; errors began to be seen as primary indications of the difficulties learners face in their L2 learning process and assisted language teachers devise subsequent strategies to help learners overcome those barriers (Mungungu, 2010). Consequently, specification, evaluation, and correction of students' errors have increasingly attracted L2 researchers and teachers' interest (Hughes & Lascara-tou, 1982; Maharjan, 2009). Error correction is also significant for L2 writing development and has been well received by L2 learners (Hyland, 1998; Hyland & Anan, 2006; Tran, 2013). Recent studies indicate that teachers' correction of errors in learners' writing assignments works to their benefits and improves their writing (Bitchener & Knoch, 2010; Karim & Nassaji, 2018, 2020). The degree of usefulness of such a correction, nevertheless, is substantially dependent on consistency and uniformity of error judgment and evaluation made by teachers (Hyland & Anan, 2006).

Error correction, in practical terms, incorporates three processes of identification, correction, and evaluation, although most studies have mainly emphasized on identification and correction, leaving out error evaluation (Hyland & Anan, 2006). Sheorey (1986) referred to error judgment and evaluation as a complex and less appreciated task, which has not sufficiently been addressed in error treatment studies. He states that research has mainly focused on error correction strategies that teachers use, and the effects theses corrections have on students' learning and the critical issue of how teachers evaluate those errors has received low profile. The main problem teachers may face in error judgment is the lack of access to guidelines or criteria that specify the degree of seriousness or gravity of those errors (Grobe & Renkl, 2007). For this reason, the debate on error judgment has leaned towards criteria development in order to avoid ambiguous and inconsistent correction practice (Hyland & Anan, 2006). Since error judgment somehow depends on shared opinions, one rational way to develop error correction criteria is to consult the stakeholders who are involved in the writing process (Vann et al., 1984). As such, researchers have been obsessed with native speakers' reactions to students' errors as evaluation criteria and strived to measure which errors seems unacceptable or interfere with comprehension from the viewpoints of native speakers (Vann et al., 1984). Later, a few studies have addressed the differences in perceptions of error gravity between native and non-native teachers. These studies revealed difference of opinions between native and non-native teachers in judgment of students' errors in writing (e.g., Hyland & Anan, 2006; Hughes & Lascaratou, 1982; James, 1977; Lee, 2009; Shi's, 2001). As pointed, work on error analysis has focused on teachers' (either native or no- native) judgments of error acceptability (Burt & Kiparsky, 1975) and learners' views regarding seriousness of errors have been ignored. Learning about English as foreign language (EFL) learners' judgments as one of the main stakeholders in academic EFL programs can be a practical step in developing priorities for EFL writing instruction. This implication is perfectly in line with conclusions made by Sharifian (2009) and Marlina (2014) who state that the English as International Language (EIL) disregards reliance on a particular variety of English and attributes English a pluricentric position where all varieties are acknowledged for international communications and intercultural relationships. In other words, from a pedagogical view of EIL, the aim of English instruction is to make EFL learners competent communicators with all kinds of world English users, and not merely with a selected group of native speakers of English or competent interlocutors like teachers, a view which holds a separate position from existing ELT trend that highlights the inner-circle varieties and disregards non-native speaker (NNS) English varieties (Sharifian, 2009). One of the probable outcomes of this movement is that teacher's focus mainly on the errors that are of greater gravity in the view of the learner and align their priorities with theirs.

Among the many existing categories of errors, lexical errors are a prevalent type of error received little attention (Hemchua & Schmitt, 2006; Kroll, 1990; Llach, 2007). To date, few if any, research has been conducted on investigating the gravity of EFL learners' written lexical errors from the viewpoint of their peers. As such, the present study attempts to fill this gap and investigates the perceptions of gravity of advanced EFL learners' written lexical errors from their peer advanced learners' perspective. Error perceptions can lay a fruitful pedagogical ground for an effective error treatment in the future (Brown, 2014). Considering the global trend to acknowledge that English is nowadays used for communication amongst NNSs, investigating the gravity of the errors made by the EFL learners from the perspective of other EFL learners can be fruitful as it can tell us about which lexical errors are perceived grave by those learners in the production of their fellow EFL learners. Thus, the following research questions will be addressed in this study:

1. What type of lexical errors do Iranian advanced EFL learners find least acceptable in their peers' written language?
2. What type of lexical errors do Iranian advanced EFL learners find least intelligible in their peers' written language?
3. Is there a correlation between ratings of acceptability and intelligibility lexical errors?

2. LITERATURE REVIEW

L2 writing teachers are commonly confronted with the problem of dealing with learners' errors. Errors; however, are common aspect of the process of language acquisition and provide teachers and researchers with valuable information regarding students' learning (Corder, 1967; Dulay & Burt, 1974; Nemser, 1971; Richards, 1971; Selinker, 1972). A distinction is usually made between errors and mistakes. While mistakes are regarded as performance-related errors in spontaneous speech or writing, which happens in the native speakers' speech as well, errors are characterized as deviations from the standards of the target language produced by non-native speakers (Lee, 1990). Errors are considered a useful pedagogical tool in the context of L2 teaching too. Gass, Behney and Plonsky (2013) state that errors are like red flags as they give the L2 instructors hints about the learners' evolving internal system and where they need help, thus enabling them to develop materials and techniques to assist learners overcome those errors.

According to Corder (1987), language learners errors are worth analyzing for three distinct reasons. First and foremost, errors and their types can be a yardstick through which the level of learners' existing L2 knowledge can be measured. Hence, lapses in the learners' competence are indicator of what they have to acquire further in their L2 learning process. Second, there is a linear relationship between learners' L2 acquisition and the errors they commit. Third, error production on the side of learners' is accompanied by error treatment on the side of teachers. Through teachers' feedback, learners would be able to discover new rules or repair the current impaired rules in their L2 system. SLA researchers have examined L2 learners' errors from a number of perspectives. The concept of error has also been linked to irritability, defined by Ludwig (1982) as "the result of the form of the message intruding upon the interlocutor's perception of the communication" (p. 275). Besides, Ludwig (1982) described irritation as a continuum ranging from an unconcerned, undistracted awareness of a communicative error to a conscious preoccupation with form. Other studies (e.g., Meyer & Lorenz, 1984; Santos, 1988; Vann et al., 1984) have regarded irritation or stigmatization as an entirely subjective criterion dependent on the outlook of the hearer or listener. Since correction of students' written

errors are generally considered as fundamental for writing development (Ferris, 2002), teachers have to come to a decision as whether or not to correct learners' errors and if they decide to do so, which types of errors they think should be corrected. One strategy for developing error correction priorities is to consider the perceptions of some audiences mainly native speakers (Kalil, 1985) and teachers (Vann et al., 1984) regarding seriousness of errors and the effect those errors may have on them.

The term "error gravity" which has been used interchangeably with other terms such as 'error perception', 'error judgments', or 'error evaluation' (Endley, 2016), is characterized as an attempt to investigate the errors which are perceived to be the most serious or distracting to readers or listeners, along with factors that exert influence on such judgments (Endley, 2016). It appears, therefore, that error gravity generally depends on the attitude of the listener or reader and cannot be judged by a universal criterion (Lee, 1990). As a research topic, error gravity came to the fore in the 1970s and 1980s but seemed to have fallen out of rigor in the 1990s. Only now, in the last few years, have researchers begun to show a renewed interest in this issue (Endley, 2016). Error gravity introduces a criterion for error correction, indicating the categories and instances of error, which need urgent pedagogical attention (Corder, 1975; Lee, 1990; Richards, 1971). Hence, by developing different taxonomies of errors, research in this area aims at identifying instructional priorities which would enable instructors working in different pedagogical settings to attend to problems arisen in this regard (Endley, 2016). Various researchers have strived to find correlations between perceptions of error gravity and one or more variables such as the raters' teaching experience (James, 1977; Oliaei & Sahragard, 2013), their area of academic specialization (Meyer & Lorenz, 1984; Roberts & Cimasko, 2008; Santos, 1988; Vann et al., 1984;), and their age (Vann et al., 1984). However, one more variable that has received considerable attention in this respect is the notion of native speakers (NSs) versus NNSs' judgments on L2 learner errors (Rao & Li, 2017). The findings of a multitude of studies (e.g., Hughes & Lascaratou, 1982; Hyland & Anan, 2006; James, 1977; Porte, 1999; Rao & Li, 2017) in this respect demonstrate that generally NNSs are less tolerant of learner errors and tend to rely on rule infringement; on the other hand, NS teachers are more lenient and prioritized intelligibility in evaluating errors.

While most studies on error judgment have focused on measuring native speakers' reactions, Khalil (1085) recommends rather than relying on native speakers' judgment, researchers are required to look for "alternative means of judging communicative effect to establish pedagogically appropriate and useful hierarchies" (p. 384). Moreover, most studies so far have focused on judgment of teachers and error perceptions and judgment of EFL learners who are the main actors in language learning process have not been taken into consideration. A successful L2 writing class requires that both parties, the teacher and learners, be actively involved in the learning and teaching process (Tran, 2013). From the viewpoint of EIL (Alsagoff et al., 2012; Matsuda, 2107; Mckay, 2018; Mckay & Brown, 2016; Sharifian, 2009), which is a new approach to teaching English, it has been extensively discussed that English is not limited to native- speaking communities since the number of people who speaks English as a second or foreign language goes far beyond that of native English speakers (e.g., Lluraa, 2017; Marlina, 2018; Schuttz, 2019). The global expansion of English has thus eventuated in revisiting how this language should be taught and conceptualized (Mckay & Brown, 2016; Nushi, et al., 2016). Accordingly, "the unprecedented global demand, use, and appropriation of EIL necessitate a profession-wide response to English language learning,

teaching, teacher education, assessment, and policy" (Selvi, 2013, p. 42). EIL takes into account the fact that the use of language in interaction is heavily contingent on mutual intelligibility between the interlocutors, the speaker's current level of expertise, and the listener's English proficiency (Mckay, 2018). Elaborating the concept of mutual intelligibility, Seidlhofer (2011) contends that intelligibility in communication is influenced by not only language skills but also perceptions of those who are being addressed. She maintains that the perception of our addressees and whether they belong to the same social or ethnical group affect our expectations in linguistic exchanges and determine the degree to which speakers comprehend each other's messages. Mckay (2018) states that EIL has been characterized as both the many varieties of English spoken globally and the use of English by NNSs. As Stern (1983) puts it, EIL leaves no room for the concept of native-like competence since the majority of EIL learners use English along with other languages and follow some certain motives for learning English that are different from those of native speakers. As such, EIL has gained momentum in academic venues and is considered as a legitimate alternative to the traditional ESL/EFL dichotomy. In view of this, investigating EFL learners' perception and judgment of errors is essential.

Lexis is one of the significant elements of written language. Learners need to make use of words accurately in written communication to get their messages through. Ellis (1994) states that the most prevalent type of errors non- native language learners make is lexical in nature. According to Llach (2011, the term lexical error is used to refer mostly to "the deviations in the learner's production of the L2 norm with regard to the use in production and reception of lexical items" (p. 75). Lexical errors are worth attending to for a number of reasons. First and foremost, lexis is one of the important aspects of successful communication, especially in the written mode. Folse (2004) holds that "with poor vocabulary, communication is constraint considerably. You can get by without grammar; you cannot get by without vocabulary" (p. 2). Second, some studies (e.g., Lennon, 1991; Meara, 1984) suggest that EFL learners are more susceptible to committing lexical errors than other types of linguistic errors and that these errors are of high frequency in L2 learners' writing (e. g., Ahn & Kang, 2015; Lee, 2017). Third, lexical errors can affect the quality of academic writing (Engber, 1995), bring about senses of intolerance and irritation in exchanges between native and non-native speakers, and are more likely to impede the flow of communication than their counterparts in syntax (Carter, 1998; Saud, 2018). Finally, possessing a good command of lexical knowledge plays a pivotal role in L2 learning as corroborated by Schmitt (2000) who asserts "lexical knowledge is central to communicative competence and to the acquisition of the second language" (p. 55).

Different taxonomies have been developed to categorize and further analyze lexical errors (e.g., Djokic, 1999; Engber, 1995; James 1998; Lennon, 1996; Warren, 1982). Drawing on the form versus content-oriented distinction, James (1998) developed a lexical error taxonomy that is comprised of two major categories namely, formal and semantic. The formal errors are divided into three major classes: (formal mis-selection, mis-formation, distortion) which in turn are accompanied by twelve subclasses. On the other hand, the semantic errors are divided into two major categories (confusion of sense relations, collocation errors), which are in turn subdivided into five classes (Figure 1. James' (1998) lexical error taxonomy).

Figure 1: James' (1998) lexical error taxonomy

I. Formal Errors

1. Formal mis-selections
    1.1. Suffix type
    1.2. Prefix type
    1.3. Vowel-based type
    1.4. Consonant-based type
2. Mis-formation
    2.1. Borrowing (L1 words)
    2.2. Coinage (inventing based on L1)
    2.3. Calque (Translation from L1)
3. Distortions
    3.1. Omissions

    3.2. Over-inclusion
    3.3. Mis-selection
    3.4. Mis-ordering
    3.5. Blending

II. Semantic Errors

1. Confusion of sense relations
    1.1. General term for specific ones
    1.2. Overtly specific term
    1.3. Inappropriate co-hyponyms
    1.4. Near synonyms
2. Collection Errors

So far, a multitude of studies have been conducted on gauging the frequency or problematizing the nature of lexical errors (e.g., Hemchua & Schmitt, 2006; Shalaby, Yahya & El- Komi, 2009; Hamadi, 2016; Saud, 2018). Relying on James' (1998) taxonomy of errors, Hemchua and Schmitt (2006) analyzed lexical errors of Thai university EFL students extracted from their compositions. The results indicated that semantics, more than the forms of words, cause communication problems for the students. For formal errors, formal mis-selection followed by distortions and for semantic errors, collocations followed by confusion of sense relations were found to be the most frequent. In a similar vein, Shalaby et al. (2009) scrutinized the types of lexical errors in the writing of Saudi college students. The results revealed that lexical semantic errors were more frequent than lexical formal errors. Besides, mis-selection of suffix type was found as the most problematic error category and lexical errors of L1 direct translation was next in frequency. The most frequent error types in the lexical semantic category were confusion of sense relations and collocation errors respectively. Hamdi (2016) also analyzed the lexical errors by Tunisian EFL learners via James' taxonomy. The findings indicated that lexical formal errors were higher in frequency in comparison with lexical semantic ones. Additionally, distortion followed by mis- formation was the most problematic and formal mis-selection were the least problematic errors. Saud (2018) also examined Saudi EFL learners' lexical errors through employing an achievement test. Employing James' comprehensive taxonomy of errors, he concluded that formal mis-selection error type was the most frequent category of formal errors and confusion of sense relations was the most frequent

among the lexical semantic errors. Vowel-based error was found to be the most problematic formal error and use of near synonym was the most problematic semantic error.

As the literature indicates, there has been a wealth of studies done on investigating the lexical errors; however, few, if any, studies have examined EFL learners' perceptions towards these errors. This is of particular importance given the fact that in the current EIL paradigm what matters in NNS-NNS communication may be different from that in native speaker-native speaker (NS-NS) or NS-NNS communication. To fill this void, the present study has taken a new step forward and aims to determine the gravity of lexical errors made by Iranian EFL learners, in terms of their acceptability and intelligibility, as perceived by their peer learners.

3. METHOD

Participants

The participants of this study consisted of 60 Iranian undergraduate students majoring in English Language and Literature at Shahid Beheshti University of Tehran, Iran. These participants, both male and female, were within the 18-22-age bracket and were selected through a purposive sampling to meet the inclusion criterion of this study, that is, they had to be advanced EFL learners. They were given a questionnaire containing eleven lexically erroneous sentences and asked to judge those sentences in terms of acceptability and intelligibility. The sentences were extracted from their peers' (fellow advanced students) IELTS writing task 2 essays and classified based on James' (1998) taxonomy of error (explained later in the article).

Instruments

IELTS Writing Task 2

The students were given an IELTS Writing Task topic to write on and their essays were examined and scored according to the standard IELTS writing exam rubric. These writing scores were then matched to the set of Common European Framework of Reference for Languages (CEFR) levels. Those whose scores fell between 7.5-8.5 (C1) were considered advanced learners (writers). To analyze and classify the lexical errors in the students' essays, James' (1998) taxonomy of lexical errors was adopted. The taxonomy is one of the most comprehensive classification systems and is based on the distinction between form- oriented and content-oriented lexical errors. The justification behind this distinction is that mental lexicon is organized in a way that it follows both formal and semantic principles. Hence, in this sense, it is either formal or semantic associations that activate the storage or access of words (Llach, 2011). James (1998) categorizes lexical errors into two major types: formal and semantic features as defined below in some detail:

Formal Errors

Formal errors are categorized into three types: (1) Formal mis-selection, (2) mis-formations, and (3) distortions. The subcategories and examples for each type are stated below:

Formal mis-selection or the incorrect word choice involve similar lexical forms and is of four major types:

The suffix type: These are lexical errors in which the roots are the same but suffixes are different (for example, competition/competitiveness)

The prefix type: These are lexical errors in which the roots are the same but prefixes are different (for example, reserve/preserve)

The vowel-based type. These are lexical errors, which contain wrong choice of vowels (for example, seat/set)

The consonant based type: These are lexical errors, which contain wrong choice of vowels (for example, save/safe).

Mis-formations are words that do not exist in the L2. The source of these errors is from the learner's L1. These errors are classified into three types:

Borrowing involves using L1 words in L2 without any change

Coinage involves learner formulating a new word from L1

Calque involves translating one word or phrase from learners' L1

Distortions refer to words that do not exist in L2. There is no transfer from L1 and the words are produced as the result of misapplication of L2. James (1998) classifies distortions into five types:

Omission (**intresting** instead of interesting)

Over inclusion (**dining** room instead of dining room)

Mis-selection (**delitous** instead of delicious)

Mis-ordering (**littel** instead of little)

Blending (**travell** instead of travel).

Semantic errors

In addition to formal errors, James highlights two types of semantic errors:

Confusion of sense relations, and

Collocation errors.

Confusion of sense errors: These are errors in which a word is used in contexts where a similar word should be used and encompasses four types of errors:

Using a superonym for a hyponym refers to using a more general term instead of a specific one.

Using a hyponym for a supersonic involves using **a** specific term instead of a general term

Using inappropriate co-hyponyms

Using a wrong near synonym.

Collocation errors: The second type of semantic errors is referred to as collocation errors are those types error in which a wrong and inappropriate word is selected to accompany another word.

Data Collection Procedure

Eleven lexically deviant sentences were randomly chosen from amongst the erroneous sentences in the students' essays. These sentences were selected in such a way that they represent different categories of lexical errors. Caution was applied to make sure that each sentence included only one error. The errors were written in bold in each sentence for easy spotting. The respondents were then asked to judge the gravity of the errors in those sentences in terms of acceptability and intelligibility on a five-point Likert scale attitude questionnaire (see the Appendix). The Likert scale ranged in an ascending order from 1 (Not acceptable/intelligible) to 5 (perfectly acceptable/intelligible). The reliability of the questionnaire calculated via Cronbach's Alpha turned out to be .82, indicating that the questionnaire enjoyed a good level of reliability. Additionally, five EFL experts who held PhD degrees in TEFL and had 10 years of teaching experience examined the content validity of the questionnaire. Five advanced EFL learners also reviewed the questionnaire. The questionnaire was revised based on the comments received and then administered to the participants to complete the questionnaire. The two elements of acceptability and intelligibility were defined in operational terms. Acceptability was defined as the degree to which the errors were assessed in language use as acceptable or unacceptable in different contexts with particular purposes and intelligibility was defined as the degree to which the error in language use could be considered as ambiguous or understandable in different contexts.

4. RESULTS

Table 1 displays the mean of students' acceptability ratings of 11 categories of errors on the 5-point scale as well as the standard deviations of their ratings. The answer to research question one can be revealed by examining the Table, that is, the learners tended to consider 'mis-ordering' as the least acceptable error type (error 7), with a mean value of 1.78. Mis-ordering belongs to the category of formal errors and refers to words that do not exist in the L2; the source of these errors often goes back to the learners' L1. The questionnaire item that contained the 'mis- ordering' error was: *I **had myself** this terrible stress, which was caused by konkor* (exam to enter Iranian university). The analysis also showed that the most acceptable error type was 'confusion of sense relations' (error 9) (M= 3.53) in the semantic-type category that included choosing a specific term (hyponym) where a more general one (superonym) was needed. The questionnaire item that contained the 'confusion of sense relations' was: *Migrating from one country to other countries isn't a new phenomenon. More skilled people leave their home **towns** in order to have better employment choices.*

Table 1: Descriptive Statistics for the Acceptability of the Eleven Lexical Errors

Errors		1	2	3	4	5	6	7	8	9	10	11
N	Valid	60	60	60	60	60	60	60	60	60	60	60
	Missing	0	0	0	0	0	0	0	0	0	0	0
Mean		3.03	2.50	3.06	3.35	2.18	2.40	1.78	2.41	3.53	3.13	2.73
Std. Deviation		1.23	1.32	1.28	1.44	1.29	1.18	1.05	1.40	1.12	1.47	1.21

Note (for Table 1.)

Error 1: Formal mis-selection/suffix type Error 2: Formal mis-selection/prefix type Error 3: Mis-formation/borrowing
Error 4: Mis-formation/coinage
Error 5: Mis-formation/calque
Error 6: Distortion/mis-selection
Error 7: Distortion/mis-ordering
Error 8: Confusion of sense relations/use of a superonym for a hyponym
Error 9: Confusion of sense relations/use of a hyponym for a superonym
Error 10: Confusion of sense relations/use of inappropriate co-hyponyms
Error 11: Collocation errors

In terms of intelligibility (the second research question), like the case in acceptability, learners found the use of specific term (hyponym) instead of a general (superonym) term as the most intelligible error (*M* = 4.13) that is included in the semantic, 'confusion of sense relations' category. Subsequently, the least intelligible error turned out to be of formal error category, that is, mis-formation, 'calque' (i.e., error 5, translation from L1) (*M* = 2.46) (see Table 2). The questionnaire item that contained the 'calque' error was: *If we **care**, celebrities are not very different from us.*

Table 2: Descriptive Statistics for the Intelligibility of the Eleven Lexical Errors

Errors		1	2	3	4	5	6	7	8	9	10	11
N	Valid	60	60	60	60	60	60	60	60	60	60	60
	Missing	0	0	0	0	0	0	0	0	0	0	0
Mean		3.80	3.88	3.83	3.88	2.46	2.85	2.48	3.68	4.13	3.73	3.41
Std. Deviation		.97	1.02	1.22	1.19	1.33	1.14	1.37	1.15	1.19	1.26	1.19

Table 3: Correlations between Acceptability and Intelligibility of the Errors

		Acceptability	Intelligibility
Acceptability	Pearson Correlation	1	.408**
	Sig. (2-tailed)		.001
	N	60	60
Intelligibility	Pearson Correlation	.408**	1
	Sig. (2-tailed)	.001	
	N	60	60

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

To answer the third research question, Pearson's correlation was used to probe the correlation between acceptability and intelligibility ratings of the students. The results indicated that there was a strong positive correlation between acceptability and intelligibility ratings of the errors by the advanced EFL learners, meaning that the more acceptable the errors, the more intelligible they were (see Table 3).

5. DISCUSSION

Previous studies suggest that lexical errors are the most prevalent category of errors in written English (e.g., Llach, 2011). Studies on error gravity have also revealed that lexical errors are judged to be more serious than their structural counterparts as they cause the greatest interference to successful communication. One prevalent strategy to develop error correction criteria is to consult the stakeholders who are involved in the writing process (Kalil, 1985; Vann et al., 1984). While earlier work on error gravity has mostly emphasized native speakers' judgments of error acceptability (Burt & Kiparsky, 1974), EFL learners' views regarding seriousness of errors have been ignored. Learning about EFL learners' judgments of typical EFL writing errors as the main stakeholders in EFL programs can improve teachers' understanding of the of errors and help them align their error correction priorities with students' and focus principally on the errors that are of greater gravity in the view of the learners. As a result, the primary purpose of this study was to gauge the gravity of written lexical errors made by Iranian EFL learners from their peer perspective; the gravity of errors was measured in terms of acceptability and intelligibility.

The results of this study revealed that the learners marked the semantic lexical errors as more acceptable and more intelligible in comparison with the formal lexical errors, implying that EFL learners face more difficulty when confronting the formal errors. This finding is in line with previous studies, which demonstrated that the participants committed more formal errors than semantic errors (Amin, 2014; Hamdi, 2016; Rezai & Davarpanah, 2019; Saud, 2018). The results

also indicated that the least acceptable type of errors was mis- ordering (at the lexical level). Mis-ordering error alters the surface structure of the sentence and may emanate from incomplete application of rules, that is, failure to fully develop a structure. According to Ridha (2012), mis-ordering is a common error amongst L2 learners. Nushi (2016) has also found out that word order errors are the likely candidates of fossilization in the Iranian EFL learners' interlanguage and EFL practitioners need to give serious pedagogical attention to teaching of this feature to help learners avoid fossilization. Furthermore, Vann et al. (1984) stated that errors such as inappropriate preposition or lack of pronoun agreement have been described in the literature as less grievous, while those errors that impede comprehension, such as word order and word choice have been considered as more grievous. It appears that to EFL learners in this study, mis-ordering error or incorrect placement of an item in a sentence is the type of error that makes the meaning of the sentence more ambiguous and incomprehensible and globally impact the communication (Dulay, Burt, & Krashen, 1982). This perception of gravity of mis-ordering errors supports findings of Burt and Kiparsky (1974) regarding the undesirable communicative effect of errors in areas like word order, which tends to have an effect on whole sentence organization.

The least intelligible type of error, on the other hand, was calque. Calque is a word-for- word translation from one language to another. It is an interlingual error (i.e., negative transfer), which falls within lexical errors (inappropriate creation) and its occurrence indicates that students make frequent use of existing terms incorrectly by amplifying their meaning content. Errors of this type lead learners to directly translate from their mother tongue because of their literal meaning. In other words, L2 learners generally commit transfer errors, as they cannot use the full cluster of semantic features of the target language lexical item.

The study further revealed that the use of hyponyms instead of superonyms in the category of 'confusion of sense relation' was the most acceptable and most intelligible error type to the respondents. Semantic confusion refers to the error made by the learners in using two words of the target language, which share semantic similarity. It is speculated that this type of error is due to underdeveloped knowledge of vocabulary, which means that learners have already had enough English vocabulary but cannot discern the exact usage of the words in the sentences. Some studies (e.g., Saud, 2018; Wells, 2013) suggest that this kind of error, which is categorized as a conceptual-type error, is a highly frequent one among EFL learners. Wells (2013) believes that confusion over meaning of words is not reminiscent of L1 transfer but the outcome of learners' association of the word with the literal meaning, that is, the learners prefer one word over another because of their similar properties. This indicates that non-native speakers of English, even those who enjoy high proficiency in English, face difficulties in finding the conceptual equivalence between terms and objects despite their sufficient grammatical proficiency to start and continue communication. The underlying justification for the acceptability and intelligibility of using hyponym instead of superonym in writing may be pertinent to the notion that in today's pedagogical and educational milieus, communicative language teaching is the dominant approach and, in that approach, concepts but not terms are emphasized. Mackay (2018) asserts that the use of language in interaction is heavily contingent on mutual intelligibility between the interlocutors, the speaker's current level of expertise, and the listener's English proficiency.

The results further verified a positive correlation between intelligibility and acceptability of errors, which implies that these two factors are interdependent, that is, an acceptable error is less likely to interfere with comprehension. This finding is in agreement with previous studies on error gravity that emphasized the significance of comprehensibility as judgment criterion and highlighted the potential of lexical errors to act as communication distracters. Hughes and

Lascaratou (1982) who studied error gravity judgment by native teachers found out that the vocabulary and spelling errors were the most grievous errors to native speakers because they hinder comprehension. Likewise, Khalil (1985) and Olsson (1973) found that semantically deviant utterances are less intelligible than grammatically deviant ones. In their error gravity studies, Santos (1988) and Dordick (1996) observed that lexical errors were judged by faculties as the most grievous errors since they make the meaning of the message ambiguous and unintelligible, leading to comprehension difficulties and accordingly disrupting communication. In other words, judgment of acceptability of errors is so dependent on the degree of unambiguousness or intelligibility of the message conveyed and when lexical errors interfered most with communication, they were marked as the most serious ones.

6. CONCLUSION

The present study was designed to determine which sentence-level errors are judged to be most serious by Iranian EFL students to help develop an error correction priority for EFL writing teachers. A survey was conducted to measure how Iranian EFL students at Shahid Beheshti University reacted to written lexical errors their classmates had made in an IELTS writing task 2. Sixty respondents judged the relative gravity of the written errors that occurred in 11 sentences. The results of the study showed that the participating students did not consider all errors as equally serious; rather, their judgments generated a hierarchy of errors. In this regard, participants marked the semantic lexical errors as more acceptable and more intelligible in comparison with the formal lexical errors. They also marked calque as the least intelligible type of error and the use of hyponyms instead of superonyms as the most acceptable and most intelligible error type. The analysis of the data also verified a positive correlation between intelligibility and acceptability of errors.

There is now little doubt that English has established itself firmly as a world language through which millions of speakers with different first languages communicate. In fact, the number of those who speak English as an L2 now far exceeds that of those who speak it as their native tongue (Clyne & Sharifian, 2008; Crystal, 2003; McKay, 2002; Nault, 2006; Rubdy & Saraceni, 2006). In our field, TEFL or TESOL, the recognition of internationality of English is evidenced in the popular use of terms such as English as a Lingua Franca (Jenkins, 2000, 2007, 2009; Mauranen & Ranta, 2009), English as an international language (EIL) (Baxter, 1980; Smith, 1983) and English as a Global Language (Crystal 2003; Gnutzmann, 1999), World Englishes (Kachru, 1985). There are now even journals (e.g., *Journal of English as a Lingua Franca* and *Journal of English as an International Language*) that exclusively focus on issues related to the rapidly-growing use of the English language worldwide. This widespread use of English for communication amongst non-native speakers with different L1 backgrounds requires us "to revisit how we define and assess language proficiency" in English (Nushi, Abulhassani, & Mojerloo, 2016); the aim of English instruction should be to make learners competent EIL communicators. Holding such a view entails moving away from the current ELT practices that highlight the inner-circle varieties at the expense of NNS English varieties. In a similar vein, Sharifian (2009) considers EIL as the language of international and intercultural relations, which "rejects the idea of any particular variety being selected as a lingua franca for international communication" (p. 2). Such a shift in attitude towards proficiency in English also means a fresh look at EFL learners' lexical use and particularly the errors they make when using the lexical items in the target language. Of the viable implications of this study can be providing the grounds for EFL teachers to address these issues and become cognizant of the principles of EIL approach through using supplementary materials and encouraging learners to be sensible of other varieties of English. Learners' exposure to and

familiarity with different varieties of English can equip them better reinforcements for international communication and greater tolerance of non- native local varieties (Rahimi & Pakzadian, 2019).

L2 learners' writings may contain lexical formal and semantic errors, which mean they have not developed adequate vocabulary to accurately express themselves, either due to language transfer or inadequacy of instruction or insufficient practices. Instructors therefore could assist them by developing activities for those areas that they need improvement. Based on judgments of L2 learners about the seriousness of lexical errors, an error-gravity scale can be developed from the most serious or grievous errors to the least serious ones to inform teachers' instructional error-evaluation practices (Chan, 2010). This hierarchy of error gravity can assist writing teachers with planning and administering writing instruction in order to fill the existing gaps and make more informed decisions regarding their priorities in treating errors. Based on the findings of present study, it can be suggested that EFL teachers should be more tolerant of semantic confusion errors and give higher priority to mis-ordering and calque errors and provide students with appropriate feedback and explicit instructions. We end this discussion by suggesting that future researchers review the pertinent literature within the realm and scope of this study and observe to what extent the priorities of the EFL teachers match those of the EFL learners in terms of error correction and determine the measures that can be taken to narrow down the prospective mismatch.

The study has some limitations. First, the study focused on errors at sentence-level, detached from context and this would have influenced on audiences' judgments. Second, the number of lexically erroneous sentences to be judged by students was small and as a result not all categories of lexical errors illustrated in James' taxonomy were incorporated. A great deal more research needs to be conducted on the EFL students' judgment, with greater number of students and greater number of error types to develop a set of reliable lexical error correction criteria. As such, future studies may be conducted to address other types of lexical errors and to include EFL students from different contexts such as school or language institutes. Third, the effect of students' background characteristics including age, academic discipline, or gender on their judgment were not taken into account so future researchers may attempt to focus on how different background features affect the participants' judgment.

References

Ahn, M., & Kang, N. (2015). An analysis of written errors of Korean adult learners of English. *Modern English Education, 16*(2), 67–89.

Alsagoff, L., McKay, S. L., Hu, G., & Renandya, W. A. (Eds.). (2012). *Principles and practices for teaching English as an international language*. Routledge.

Amin, B. (2014). Lexical error in writing English words made by students of the junior high school. *Exposure, 3*(1), 107–133.

Baxter, J. (1980). Interactive listening. In L. Smith (Ed.), *Readings in English as an international language* (pp. 78–94). Pergamon Press.

Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten-month investigation. *Applied Linguistics, 31*(2), 193–214. https://doi.org/10.1093/applin/amp016

Burt, M. K. (1975). Error analysis in the adult EFL classroom. *TESOL Quarterly, 9*(1), 53–63.

Burt, M. K., & Kiparsky, C. (1974). Global and local mistakes. In J. H. Schumann & N. Stenson (Eds.), *New frontiers in second language learning* (pp. 71–79). Newbury House Publishers.

Candlin, C. N. (1975). *The communicative teaching of English*. Longman.

Carter, R. (1998). *Vocabulary: Applied linguistics perspectives* (2nd ed.). Routledge.

Chan, A. Y. (2010). Toward a taxonomy of written errors: Investigation into the written errors of Hong Kong Cantonese ESL learners. *TESOL Quarterly, 44*(2), 295–319. https://doi.org/10.5054/tq.2010.219945.

Clyne, M. G., & Sharifian, F. (2008). English as an international language: Challenges and possibilities. *Australian Review of Applied Linguistics, 31*(3), 1–16. https://doi.org/10.2104/aral0831

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*(4), 161–170. https://doi.org/10.1515/iral.1967.5.1-4.161

Corder, S. P. (1975). Applied linguistics and language teaching. *Papers in Applied Linguistics, 2*, 1–15.

Crystal, D. (2003). *English as a global language*. Cambridge University Press.

Crystal, D., & Varley, R. (2013). *Introduction to language pathology*. John Wiley & Sons.

Davies, E. E. (1983). Error evaluation: The importance of viewpoint. *ELT Journal, 37*(4), 304–311. https://doi.org/10.1093/elt/37.4.304.

Djokic, D. (1999). Lexical errors in L2 learning and communication. *Rassegna Italiana di Linguistica Aplicata, 31*, 123–135.

Dordick, M. (1996). Testing for a hierarchy of the communicative interference value of ESL errors. *System, 24*(3), 299–308. https://doi.org/10.1016/0346-251X(96)00023-8

Dulay, H. C., & Burt, M. K. (1974). Errors and strategies in child second language acquisition. *TESOL Quarterly, 8*(2), 129–136. https://doi.org/10.2307/3585536

Dulay, H., Burt, M., & Krashen, S. D. (1982). *Language two*. Oxford University Press.

Endley, M. J. (2016). UAEU faculty perceptions of error gravity: A preliminary investigation. In A. Elsheikh, J. Riddlebarger, S. Abu-Rmaileh, & S. Al Alami (Eds.), *Proceedings of the 21st TESOL Arabia conference: Theory, practice, innovation – Teaching and learning in the digital world* (pp. 85-102). TESOL Arabia Publications.

Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing, 4*, 139–155.

Ferris, D. (2002). *Treatment of error in second language student writing*. The University of Michigan Press.

Folse, K. (2004). Myths about teaching and learning second language vocabulary: What research says? *TESL Reporter, 37*(2), 1–13.

Gass, S., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course* (4th ed.). Routledge.

Gnutzmann, C. (Ed.). (1999). *Teaching English as a global language: Native and non-native perspectives*. Stauffenburg.

Grobe, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17*(6), 612–634. https://doi.org/10.1016/j.learninstruc.2007.09.007

Hamdi, S. (2016). An analysis of lexical errors in the English compositions of EFL Tunisian learners. *International Journal of Humanities and Cultural Studies, 2*(4), 643–652.

Hemchua, S., & Schmitt, N. (2006). An analysis of lexical errors in the English compositions of Thai students. *Prospect, 21*, 3–25.

Hughes, A., & Lascaratou, C. (1982). Competing criteria for error gravity. *ELT Journal, 36*(3), 175–182. https://doi.org/10.1093/elt/36.3.175

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing, 7*(3), 255–286. https://doi.org/10.1016/S1060-3743(98)90017-0

Hyland, K., & Anan, E. (2006). Teachers' perceptions of error: The effects of first language and experience. *System, 34*(4), 509–519. https://doi.org/10.1016/j.system.2006.09.001

James, C. (1977). Judgments of error gravities. *English Language Teaching Journal, 31*(2), 116–124. https://doi.org/10.1093/elt/xx.2.116

James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Longman.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford University Press.

Jenkins, J. (2009). English as a lingua franca: Interpretations and attitudes. *World Englishes, 28*(2), 200–207. https://doi.org/10.1111/j.1467-971X.2009.01577.x

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.

Karim, K., & Nassaji, H. (2018). The revision and transfer effects of direct and indirect comprehensive corrective feedback on English-as-a-second-language (ESL) students' writing. *Language Teaching Research, 23*(4), 487-507. https://doi.org/10.1177/1362168818802469

Karim, K., & Nassaji, H. (2020). The effects of written corrective feedback: A critical synthesis of past and present research. *Instructed Second Language Acquisition, 3*(1), 28–52.

Khalil, K. (1985). Communicative error evaluation: Native speakers' evaluation and interpretation of written errors of Arab EFL learners. *TESOL Quarterly, 19*(2), 335–352. https://doi.org/10.2307/3587011

Larsen-Freeman, D., & Anderson, M. (2013). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.

Lennon, P. (1991). Error and the very advanced learner. *International Review of Applied Linguistics, 29*(1), 31–44. https://doi.org/10.1515/iral.1991.29.1.31

Lennon, P. (1996). Getting 'easy' verbs wrong at the advanced level. *International Review of Applied Linguistics, 34*(1), 23–36. https://doi.org/10.1515/iral.1996.34.1.23

Lee, N. (1990). Notions of 'error' and appropriate corrective treatment. *Hong Kong Papers in Linguistics and Language Teaching, 13*, 55–69.

Lee, S. (2017). Raising EFL learners' awareness of L2 lexical errors and correct usage: A dual coding approach. *English Teaching, 72*(2), 29–50.

Llach, M. D. P. A. (2007). Lexical errors as writing quality predictors. *Studia Linguistica, 61*(1), 1–19. https://doi.org/10.1111/j.1467-9582.2007.00132.x.

Llach, M. D. P. A. (2011). *Lexical errors and accuracy in foreign language writing*. Multilingual Matters.

Ludwig, J. (1982). Native-speaker judgments of second language learners' efforts at communication: A review. *The Modern Language Journal, 66*(3), 274–283. https://doi.org/10.1111/j.1540-4781.1982.tb06972.x

Maharjan, L. B. (2009). Learners' errors and their evaluation. *Journal of NELTA, 14*(1), 71–81. https://doi.org/10.3126/nelta.v14i1.3060

Marlina, R. (2018). Revisiting the pedagogy of English as an international language. *RELC Journal, 49*(1), 3–18. https://doi.org/10.1177/0033688217753844

Matsuda, A. (2017). *Preparing teachers to teach English as an international language*. Multilingual Matters.

Mauranen, A., & Ranta, E. (2009). *English as a lingua franca: Studies and findings*. Cambridge Scholars Publishing.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, & A. Howatt (Eds.), *Interlanguage* (pp. 225–235). Edinburgh University Press.

McKay, S. L. (2002). *Teaching English as an international language*. Oxford University Press.

McKay, S. L. (2018). English as an international language: What it is and what it means for pedagogy. *RELC Journal, 49*(1), 9–23. https://doi.org/10.1177/0033688218760359.

McKay, S. L., & Brown, J. D. (2016). *Teaching and assessing EIL in local contexts around the world*. Routledge.

Mungungu, S. S. (2010). *Error analysis: Investigating the writing of ESL Namibian learners* (Doctoral dissertation). University of Pretoria.

Nault, D. (2006). Going global: Rethinking culture teaching in ELT contexts. *Language, Culture and Curriculum, 19*(3), 314–328. https://doi.org/10.2167/lcc337.0

Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics in Language Teaching, 9*(2), 115–124. https://doi.org/10.1515/iral.1971.9.2.115

Nushi, M. (2016). On the role of L1 markedness and L2 input robustness in determining potentially fossilizable language forms in Iranian EFL learners' writing. *Journal of Language and Linguistic Studies, 12*(2), 66–86.

Nushi, M., Abolhassani, Z., & Mojerloo, N. (2016). An interview with Professor Farzad Sharifian. *Asian-Pacific Journal of Second and Foreign Language Education, 1*(1), 1–10. https://doi.org/10.1186/s40862-016-0011-x

Oliaei, S., & Sahragard, R. (2013). Error treatment by experienced and inexperienced Iranian EFL teachers of writing. *Educational Research International, 2*(1), 55–75.

Olsson, M. (1973). The effects of different types of errors in the communication situation. In J. Svartvik (Ed.), *Errata: Papers in error analysis* (pp. 153–160). GWE Gleerup.

Porte, G. (1999). Where to draw the red line: Error toleration of native and non-native EFL faculty. *Foreign Language Annals, 32*(4), 426–434. https://doi.org/10.1111/j.1944-9720.1999.tb00872.x

Rahimi, M., & Pakzadian, M. (2019). English Learners' Attitudes toward English as an International Language: A Qualitative Inquiry. *i-Manager's Journal on English Language Teaching, 9*(4), 28–38.

Rao, Z., & Li, X. (2017). Native and non-native teachers' perceptions of error gravity: The effects of cultural and educational factors. *The Asia-Pacific Education Researcher, 26*(1–2), 51–59. https://doi.org/10.1007/s40299-017-0323-y

Rezai, M. J., & Davarpanah, F. (2019). The study of formal and semantic errors of lexis by Persian EFL learners. *International Journal of Cognitive and Language Sciences, 13*(1), 41–47.

Richards, J. (1971). Error analysis and second language strategies. *Language Sciences, 17*, 12–22.

Ridha, N. S. A. (2012). The effect of EFL learners' mother tongue on their writings in English: An error analysis study. *Journal of the College of Arts, 60*, 22–45.

Roberts, F., & Cimasko, T. (2008). Evaluating ESL: Making sense of university professors' responses to second language writing. *Journal of Second Language Writing, 17*, 125–143. https://doi.org/10.1016/j.jslw.2007.07.004

Rubdy, R., & Saraceni, M. (2006). *English in the world: Global rules, global roles*. Continuum.

Santos, T. (1988). Professors' reactions to the academic writing of non-native speaking students. *TESOL Quarterly, 22*(1), 69–90. https://doi.org/10.2307/3587062

Saud, W. I. (2018). Lexical errors of third year undergraduate students. *English Language Teaching, 11*(11), 161–168. https://doi.org/10.5539/elt.v11n11p161

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schutz, L. M. (2019). The cultural politics of English as an international language. *International Multilingual Research Journal, 13*(1), 70–72. https://doi.org/10.1080/19313152.2019.1557729.

Seidlhofer, B. (2011). Conceptualizing 'English' for a multilingual Europe. In A. Hu & M. Byram (Eds.), *English in Europe today: Sociocultural and educational perspectives* (pp. 133–146). Cambridge Scholars Publishing.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching, 10*(1–4), 209–232. https://doi.org/10.1515/iral.1972.10.1-4.209

Selvi, A. F. (2013). Towards EIL teacher education: Exploring challenges and potentials of MATESOL programs in the United States. In N. T. Zacharias & C. Manara (Eds.), *Contextualizing the pedagogy of English as an international language: Issues and tensions* (pp. 42–58). Cambridge Scholars Publishing.

Shalaby, N. A., Yahya, N., & El-Komi, M. (2009). Analysis of lexical errors in Saudi college students' compositions. *Journal of the Saudi Association of Languages and Translation, 2*(3), 65–92.

Sharifian, F. (Ed.). (2009). *English as an International Language: Perspectives and pedagogical issues*. Multilingual Matters.

Sheorey, R. (1986). Error perceptions of native-speaking and non-native-speaking teachers of ESL. *ELT Journal, 40*(4), 306–312. https://doi.org/10.1093/elt/40.4.306.

Smith, L. (Ed.). (1983). *Readings in English as an international language*. Pergamon Press.

Stern, H. H. (1983). *Fundamental concepts of language teaching*. Oxford University Press.

Tran, T. H. (2013). Approaches to treating student written errors. Paper presented at MIDTESOL 2013, Lawrence, Kansas.

Vann, R. J., Meyer, D. E., & Lorenz, F. O. (1984). Error gravity: A study of faculty opinion of ESL errors. *TESOL Quarterly, 18*(3), 427–440. https://doi.org/10.2307/3586715

Warren, B. (1982). Common types of lexical errors among Swedish learners of English. *Moderna Språk, 76*(3), 209–228.

Wells, N. (2013). Lexical error analysis of advanced language learners' writings (Unpublished master's thesis). The University of Wisconsin–River Falls.

Yang, Z., Wang, M., Cheng, H., Liu, S., Liu, L., & Chan, T. (2016). The effects of learning from correct and erroneous examples in individual and collaborative settings. *The Asia-Pacific Education Researcher, 25*(2), 219–227. https://doi.org/10.1007/s40299-015-0265-0

Appendix

Dear respondent, the questionnaire below aims to investigate the gravity of the different lexical errors from your perspective in two respects, namely **acceptability** (how (un)acceptable the error is in the sample sentence) and **intelligibility** (how (un)intelligible the error is in the sample sentence). The errors need to be rated on the five-point Likert scale, with:

1 being (Not acceptable/ intelligible),
2 (May be acceptable/intelligible),
3 (Acceptable/intelligible),
4 (Highly acceptable/ intelligible), and
5 (Perfectly acceptable/intelligible).

Please do not leave any of the sentences unanswered. Please highlight the circle or the number of your choice. Thank you so much for your cooperation.

#	Errors	Acceptability					Intelligibility				
1.	Some resources are not available online or in E-books forms.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
2.	It may lead to the unsatisfaction of both parents and language institutions.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
3.	They are wasting an extensive amount of time behind their computers, laptops, tablets, and cell phones.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
4.	Maybe the reasons for their success is more effort, motivation, providence, and self-belief.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
5.	If we care, celebrities are not very different from us.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
6.	There is an argument which says that libraries should be locked because better alternatives are available.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
7.	I had myself this terrible stress which was caused by Konkor (university entrance exam in Iran).	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
8.	I can have this chance to expand my vision about the Britain's language, society and people.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
9.	Migrating from one country to other countries isn't a new phenomenon. More skilled people leave their home towns in order to have better employment choices.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
10.	We should provide our students with language labs and electrical (instead	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

#	Errors	Acceptability					Intelligibility				
	of electronic) devices that ease the process of teaching and learning.										
11.	It could be avoided by arranging restrictions on the use of technology, both at home and at work.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Peer Reviewing in Applied Linguistics: Reviewers' Perceptions

Vali Mohammadi[1], Mohammad Amini Farsani[2,*] iD, Rana Nazmi[3]

[1] University of Mohaghegh Ardabili, Ardabil, Iran
[2] Iran University of Science and Technology, Tehran, Iran
[3] Imam Khomeini International University, Ghazvin, Iran

Abstract

Peer review is carried out in academic journal boards in somewhat different ways to serve the purposes of a particular journal. Through the peer review process, reviewers in academic journals scrutinize and deeply analyze the quality of academic works before the publication. As an 'occluded' genre (Swales, 1996), getting access to the content of peer reviews in journals is too difficult. To shed light on the process of peer review, we investigated the reviewers' perceptions and understandings of peer review in Applied Linguistics journals published in Iran. To this end, we developed an open-ended questionnaire and sent out it to the editorial board reviewers of Iranian certified journals active in publishing on different aspects of applied linguistics. Sixteen reviewers participated in the study by filling in the questionnaire and returning it back. The collected data were analyzed through thematic qualitative data. The results of the study indicate that the reviewers are all active agents in reviewing the manuscript and consider both conceptual, methodological, and mechanics of writing. The implications and recommendations are discussed in light of the findings.

1. INTRODUCTION

There has always been an interminable competition among the scientific institutions all over the world in terms of their academic research productions (Paltridge, 2017), and the publication of journal articles has been the harbinger of these scientific productions. Journal articles have to possess high-quality standards to be considered as a reliable academic output (Nygaard, 2015), and here the role of peer review process emerges. According to Sposato et al. (2014), a better understanding of the peer-review process could enhance the probability of publishing high-quality research. Furthermore, the 2010 Center for Studies in Higher Education Report highlights the fundamental role of the peer review in the academy (Harley, Acord, Earl-Novell, Lawrence, & Judso, 2010). Moreover, Bunner and Larsen (2012) acclaim that the peer review process has been a backbone of the scientific production for decades, rendering multiple functions such as enhancing the quality, facility, and appropriateness of the manuscripts, filtering out flawed research, and realizing a fair and unbiased assessment of a manuscript. Besides, most of the

* **Corresponding Author:** Mohammad Amini Farsani, Department of Foreign Languages, Iran University of Science and Technology, Tehran, Iran. **Email:** mohammad_farsani@yahoo.com

bibliographic databases (e.g., Scopus, ISI, JCR, SJR) consider running peer review as a basic criterion for selecting the scientific journals (Tan, Cai, Zhou, & Zhang, 2019).

Based on Committee on Publication Ethics (2012), peer review in all its forms serves a crucial function in assuring the integrity of a scholarly research; besides, the reputation of academic journals is mostly influenced by the peer review process (Paltridge, 2017). Chowdhry (2015) believes that the thought underlying the peer review process is that the flaws of a work could be better detected by a group of people other than the authors themselves and the assessments of the work would be more neutral and unbiased. Further, he assumes:

Peer review utilizes self-governance and the anonymity of the reviewers (referees) so as to discourage cronyism (i.e. bias shown to family and friends) and obtain an unbiased report. The reviewers are not selected from amongst the close colleagues/relatives/friends of the author. (p. 329) Hames (2012) accentuates that "Peer review in scholarly publishing is the process by which research output is subjected to scrutiny and critical assessment by individuals who are experts in those areas" (p. 20). Putting it in other words, Smith (2006) defines peer review as "something to do with a grant application or a paper being scrutinized by a third party - who is neither the author nor the person making a judgment on whether a grant should be given or a paper published" (p. 178). Also, Sciortino and Siemens (2013) believe that peer review "is a gate-keeper of the accepted body of scientific knowledge" (p. 225). Likewise, Trevino (2008) agrees that "peer review is an essential professional value and a duty to the profession" (p. 8). According to Herbert, Marsh, and Ball (1989), the peer review system fulfils four goals: (a) the selection of articles to be published in academic journals, (b) grant proposals to be funded, (c) individuals to be promoted, and (d) theses to be accepted as the requirements for higher degrees.

Taking the history of the peer review process into account, Schuhmann (2008) declares that the editors solely made the decisions on the rejection or publication of papers around the turn of last century and extensive review was not that much rampant; in other words, the peer review process was a private issue (Paltridge, 2017). The enforcement of peer review in academic setting leads back to approximately 300 years ago (Hames, 2012; Spier, 2002); in other terms, Leopold (2014) remarks that the peer review process roots in the 18th century. Based on Mulligan, Hall, and Raphael (2013), the Philosophical Transactions of the Royal Society of London was the first journal to perform the peer review process in assessing the academic staff. Boggs (2009) and Paltridge (2017) state that the peer review process was approved and recognized only in the late-20th century.

When an article is submitted to a journal, the editor might reject it at the initial stage (i.e., 'desk reject') or send it out for revisions by the peers. The peer reviewers could suggest publishing it as it is, they could require the authors to make some nominal corrections, or they could ask them to make major corrections. Also, Kumar, Rafiq, and Imam (2011) clarify that the typical publication process of journal submission is divided into three stages: the first stage is called pre-review (i.e., screening stage) in which the editor analyzes the article in terms of appropriateness of the subject and other generic features. The second stage called reviewing or negotiation loop where decisions on accepting, rejecting, minor corrections, major changes, or a combination of both are made; and post- review stage which includes processes for publication. The reviewers may be members of the journal's editorial advisory board or experts in the field published in the relevant domains.

If the reviewer is innominate for the author; that is, the author does not know who has reviewed the article, which is called a blind review (i.e., single-blind review). If the reviewers also do not know authors' identity either, it is called a double-blind review. If the author and the reviewers are aware of each other's identity, it is called an open peer review. For instance, an open peer review

is applied by Journals such as the British Medical Journal and Atmospheric Chemistry and Physics (Nature, 2006). Chowdhry (2015) acclaims:

In single-blind review, the reviewer identity is hidden to encourage unbiased comments, while in double-blind review, the author's identity is masked from reviewers to shield against forms of social bias. Further, an "open peer" review journal may employ a "third" party; i.e., someone who is neither affiliated directly with the reviewing entity nor associated with author being reviewed (p. 329).

Mandal, Giri, and Parija (2012) assume that the peer reviewers must be fully informed on their responsibilities as a peer reviewer. Not only must they be expert and knowledgeable in the concerned field, but also they must be vigilant on ethical aspects of research and report any encountered research malpractices or violation of ethics. Moreover, Sciortino and Siemens (2013) assert that the editors and potential peer reviewers must be transparent regarding their fields of interests and any conflicts of interest must be clarified initially. Besides, the reviewers must be dedicated and allocate adequate time and energy to commit a fair and reliable opinion. Guthrie, Parker, and Dumay (2015) express that one of the concomitant bases needed for the peer review process is trust and responsibility; however, peer reviewers are assigned to the role without appropriate guidance and awareness.

Based on Trevino (2008), devoting sufficient time and expertise while reviewing can lead to a great learning experience, as it exposes the reviewers to new thoughts and reflection, ideologies, literature, references, and data collection and analysis procedures and ultimately sparks new ideas for the future research. Similarly, Qing, Lifang, and Xiaochuan (2018) remark that through assembling of expertise of different experts, the peer review process can cover the editors' weaknesses in some specific aspects of knowledge and consequently increase the quality of academic journals. Analogous to other processes, the peer review process is also susceptible to several criticisms. Bornmann and Mungra (2011) assume that the peer review process undergoes several challenges including "its reliability and fairness; its standards, idiosyncratic and biased reviewer comments; its openness to innovation; timeliness of feedback and decisions; labour time and cost; reviewer workload; and detection of fraud and misconduct" (p.165). Besides, Paltridge (2017) suggests that the peer review process is slow, expensive, subjective, biased, and open to abuse. Hadi (2016) stipulates that because of the time- consuming nature of the peer review process and for the sake of accelerating the process, some editors may be required to recommend names of the suitable peer reviewers during manuscript submission. Some authors create fake email addresses managed by themselves; then, commit positive reports on their own manuscripts and finally get them accepted.

All these critics try to lead the peer review process to the level of constituting high- quality review. The overall thoughts of some of the reviewers cooperated in Human Resource Development Quarterly (2013) are expressed as following. Anderson suggests three characteristics for a qualified review: "balance critique with developmental intent, be open to difference, and provide feedback on a "top-to-bottom" basis" (p. 419). Furthermore, Werner accentuates that a qualified review should be honest, respectful, developmental, and timely. While Gubbins puts a great emphasis on the clarity of terms and contextual contribution as the factors leading to high quality, Lunn proves the two variables of the significance of the contribution to the field and (b) the extent to which the results of the study support the conclusions that the authors have made.

There exist a number of works examining the peer review process. For instance, Atjonen (2018) worked on the author experiences of the developmental feedback during the peer review process. The results approved the positive effects of peer-review process in terms of improving the quality

of the articles; however, the developmental feedback given to the authors needed to be emphasized more specifically. Kumar, Rafiq, and Imam (2011) focused on the main negotiation processes between the authors of articles and reviewers at the peer-reviewing stage. The results showed that the negotiations helped authors enhance the overall quality, clarity, and readability of their manuscripts.

Schwartz and Zamboanga (2015) explored ways to improve the peer review process. Mainly they scrutinized the editors' role in selecting the reviewers, adjusting their own impressions of the manuscript with the reviewers' feedback, and committing a fair and equitable editorial decision. Tan, Cai, Zhou, and Zhang (2019) investigated the relationship between the number of submissions and the overall standard of academic journals within a similar discipline. Bunner and Larson (2012) examined two online surveys, one for authors and the other for Editorial Board members to assess their perspectives on the quality and timeliness of peer review. The results revealed that perceptions of review quality among editorial board members and authors were similar, however, editorial board members were significantly more likely to rate reviewers as fair and unbiased.

Paltridge (2017) examined a study project by analyzing the reports written on submissions to the peer-reviewed journal 'English for Specific Purposes'. Reviewers also filled in a questionnaire that asked about their experience in doing peer reviews, how they had learned to write reviewers' reports, and the issues they faced in writing them. The results showed that over half of the reviewers had learned to do reviews by reading reviews of their own submissions to peer-reviewed journals. Others learnt to write reviews by just doing them, that is, by practice. The most challenging aspect for reviewers was writing reviewers' reports that were critical but still constructive. Tite and Schroter (2007) carried out a survey of peer reviewers from five biomedical journals to assess why reviewers accept or decline to review and their opinions on reviewers' incentives. Based on the results of this study, contribution of the paper to the subject area, relevance of topic to own work, and opportunity to learn something new were the main factors in deciding to accept a paper to review.

In line with all these studies, in this qualitative study, we try to contribute to the body of the knowledge on peer review through the peer reviewers' perceptions in an EFL context. Their experiences, preferences, opinions, and knowledge of reviewing are investigated. The following research questions are addressed:

1- How do EFL Iranian editorial members conceive of peer- reviewing process?
2- What are the most frequent challenges the reviewers encountered with in writing a review?
3- What are the criteria for accepting/rejecting the manuscripts?

2. METHOD

Design

In the current study, we adhered to a multiple case study design. According to this design, as Johnson and Christensen (2019) assert, it provides detailed investigation of the cases (i.e., reviewers in this study) and their perceptions of peer- reviewing process. Furthermore, the participants' perceptions were compared with each other for exploring their similar and different views toward peer-reviewing practice. As for a qualitative sampling scheme, we followed a criterion sampling strategy (Johnson & Christensen, 2019), and we gathered data from Iranian editorial board members of applied linguistics journals. It is claimed that this specific group of respondents has similar knowledge and skills in reviewing the manuscript. We sent 50 invitations to the members but 16 were agreed to cooperate with us. Of the total 16 respondents, 75% (n = 12) were male, and 25% were female (n= 4) who committed their answers within two weeks. We gave

the reviewers a consent letter at the beginning of the interview sessions. This sample size was enough to reach data saturation (see Patton, 1990)

Procedures

In order to explore what EFL reviewers had practiced and how they were able to apply their knowledge in reviewing articles, we developed an open-ended questionnaire, as one of most commonly used instruments in qualitative research (see Johnson & Christenson, 2019), based on the following steps. First, we conducted semi-instructed interviews with three experienced reviewers to develop the questions. Then, we added some items based on the related studies. To warrant the content validity of the items, two faculty members, who were all reviewers in different journals, were asked to comment on the items. The finalized version of the open-ended questionnaire was checked in terms of format, content, and appearance (Appendix I). The newly developed questionnaire was the main data gathering procedure, which was sent by an email to the editorial board members in journals. The data was gathered from Iranian editorial board members of science and research journals.

In this online survey, the e-questionnaire was sent by email to members of editorial boards of science and research journals, but because of the mass of their issues, just 16 reviewers cooperated in this study. Of the total 16 respondents, 75% (n = 12) were male, and 25% were female (n = 4) who committed their answers within two weeks.

3. RESULTS

In this section, the answers of the reviewers to the questions are analyzed. The answers to the questions 1 to 4 are categorized in Table 1.

According to Table 1, the respondents were all active in reviewing the articles, with a mean of 13.18 years of experience. More specifically, on average, they reviewed 15.5 articles in a year. Furthermore, they served 4.25 journals and spent approximately 5 hours for reviewing the articles.

Considering the question "What do you find most challenging about writing a review?" two major themes emerged under this category. They proposed two aspects: (a) those who considered challenges as one dimension (i.e., mono- aspectual reviewers), (b) those who considered challenges as Multidimensional (i.e., gestalt camp/multi-aspectual reviewers).

Mono-aspectual reviewers (n=10)

The reviewers in this category mentioned the following subthemes: (a) lack of academic literacy (n=3), (b) data analysis and discussion (n=2), (c) how to tone down criticisms (n=2), (d) making a final decision (n=1), (e) lack of novelty in articles (n=1), and (f) lack of comprehensiveness in reviewing the literature (n=1). These reviewers mentioned one of the above themes as the most challenging part of writing reviews. For example, one of the reviewers asserted that The most challenging part is when an author is not familiar with the ways a good research paper should be organized and written", at the same time another respondent was more sensitive towards technicalities of research: "Evaluation of the more scientific and technical issues is the most challenging part of reviewing task.

Table 1: A Profile of Respondents' Activity in Reviewing the Assigned Articles

IDs	No. of reviews	Editorial boards	Experiencing in reviewing (years)	Time spent on a review
1	10	10	15	5
2	5	2	10	3
3	10	7	19	1
4	30	3	10	4
5	10	4	19	15
6	8	1	8	2
7	20	5	20	9
8	20	3	10	6
9	10	1	11	3
10	5	3	10	5
11	20	7	12	3
12	15	5	10	6
13	20	5	20	7
14	20	7	12	3
15	40	3	15	4
16	5	2	10	2
Mean	15.5	4.25	13.18	4.875

Gestalt reviewers/multi-aspectual reviewers (n=6)

In terms of the most challenging aspects of the reviewing task, they expressed a multi- aspectual list of issues simultaneously. For example, one of the reviewers claimed that You need to take both form and content into account at the same time.

Moreover, two other reviewers declared that All sections of the articles are challenging from the introduction to the conclusion, even to references"; "Spotting where the research is and where it says it is challenging. And, reviewing the 'discussion' 'interpretation' of findings/results needs much work. You need to examine the research identity deeply. Third, and for some, the choice of statistical procedures by the authors… might be a source of challenge and in a good number of occasions, confusion… Why this and why not that? What if a competing alternative were used?

Taking the sixth question into account, "What do you find most straightforward about writing a review?", we found three common themes based on the respondents' attitudes. The most frequent one was related to "checking whether authors follow the guidelines" (n=5, 31%). For example, two of the reviewers answered the question by saying that when the writers have followed the exact format of the journal selected and its guidelines to prepare a qualified article" and "Checking whether authors followed the guidelines proposed by journals is the most straightforward task we can do.

The second frequent theme was related to "evaluating research methodology" (n=4, 25%). Other aspects as "Spotting the shortcomings", "spotting the quality of writing", "evaluating the persuasiveness of arguments in articles", "checking references", and "making decisions" were also mentioned in interviewing sessions.

When they were asked "how did you learn peer-reviewing?" a majority of the respondents asserted that they learned peer-reviewing via field experiencing (n=8, 50%), looking at previous reviews (n=4, 25%), trial and errors (n=2, 12.5%), and looking at journal guidelines (n=2, 12. 5%), respectively (see Table 2). This signifies that the most frequent way to learn peer reviewing was related to field experience, while the least frequent patterns were related to trials and errors and looking at journal guidelines.

Table 2: Quantification of Learning Ways of Peer Reviewing

Learn to write a manuscript review	N	Percentage
Field experience	8	50
Looking at previous reviews	4	25
Trial and error	2	12.5
Looking at journal guidelines	2	12.5
Total	16	100

In terms of the reasons for rejecting/accepting a paper, five categories are implied based on the reviewers' responses. They acclaimed that they reject the assigned papers based on the following reasons: (a) Lack of being warrantied: This category which was the most common among the reasons (n=7, 43%), contained any criticisms made by a reviewer on inappropriate, missing, and inadequate evidence. (b) Lack of novelty: This subtheme which was the second most prevalent reason (n=5, 31%), was reflected in the assertions of two respondents:

I usually reject a paper if it does not have the necessary scientific merit.
I reject a paper if it doesn't add anything to the existing knowledge.

The categories of (c) Lack of sound methodology (n=4, 25%), (d) Lack of good language (n=3, 19%), and (e) Lack of transparency were considered as other reasons for not accepting a paper. When it comes to the acceptance of a manuscript, it can be inferred from the categories above that if a paper wants to be accepted, the following criteria should be regarded: supportiveness, novelty, sound methodology, transparency, and good language.

Regarding the characteristics of good research, we concluded nine key factors from the data. A majority of the respondents within this category characterized qualified research as: 'being rigorous and methodological', 'developing well-defined problems', 'being a systematic and organized inquiry', 'well- written', 'having an adequate reporting of literature review', 'being ethical and honest', 'being interpreted well', and 'possessing novelty'. For instance, two of the respondents, respectively, noted the primacy of the strict standards of research in the following words:

One that bears added value to the theory and/or practice of ELT, has adopted a sound methodology, includes an adequate description in each of the sections of the paper (introduction, literature review, method, results, discussion, and conclusion), is well- written, and of course does not contain plagiarism.

It should have a theoretical framework or a model; good and appropriate language; comprehensive review of literature; well- elaborated method and design; well- discussed and well-documented results; and finally, an overall conclusion section to elucidate the contribution of the results of the paper to a wide range of audience and the field under investigation.

The other two respondents considered more specific characteristics of good research:

It must be innovative in the sense that it raises important issues which are related to our field by adding something to our knowledge.

It must show a true commitment of the writers in the collection and analyses of the results supported by rich discussion.

Creative question and a new perspective on the topic should be considered. Besides a sound methodology, a persuasive statement of the problem and then a significant contribution to the current state of the topic under study would be most convincing.

Regarding the last question, "When do you refer a manuscript to another reviewer?" all the respondents asserted that when they do not have the necessary expertise in the field of a paper, they suggest an alternative reviewer. For instance, they declared the followings:

When I am not an expert in the field or I want to help the writers of the article If the manuscript is not within my expertise.

When the domain of the research is out of my expertise.

When I am not sure about my knowledge of the literature in contrast, two of the respondents did not suggest an alternative reviewer.

4. DISCUSSION

The peer review process is considered a filter for scientific research productions that improve the standard of academic journals; in other words, it is central to academic publishing. As peer review guarantees the quality of the work being considered for publication (Hames, 2012), it goes without saying that peer reviewing is here to stay because all quality journals have one form of such an activity for the researchers to go with in order to publish in such academic journals (Bush, 2016).

With the rapid growth of scholarly publishing and the importance of academic writing, the peer review system has gained a fundamental role. This study tried to shed light on the peer review process from the peer reviewers' perspective. As clarified earlier in the result section, succinct responses, based on the reviewers' point of view, it is concluded that an article should be well-established, systematic, and well- designed. The research questions should be defined thoroughly and the coherence and cohesion should be regarded. The researchers should also support their research with a suitable amount of literature review and commit a sound and delightful design encouraging the reader to read the rest of the research. The results of this study elucidated that the peer review system remains useful for ensuring the quality of articles. In Sarker's (2015) words, "it is important for aspiring authors to be aware of the priorities and preferences of the audience, including the editors and reviewers" (p. 201) if they want to create a successful authorship portfolio for themselves.

The findings of this study complement the results of Bornmann and Mungra (2011). They found that the underlying theory, design and structure of the study, and the concept of the study are of high importance for the reviewers for accepting or rejecting a paper. Therefore, the implication for the novice researchers is that they have to strictly stick to the foundational fundamentals of sound research in order for their work to be publishable. In line with what Samraj (2016) argues, the take-home message for the instructors of academic writing is to encourage the graduate students to keep on (re)working on different drafts their paper so as to make it meet the paper and publication standards required by the journal.

Furthermore, Herbert, Marsh, and Ball (1989) found that a qualified study should possess a good research method and a sound writing style based on the reviewers' perspectives. Moreover, the findings of the Mulligan, Hall, and Raphael's (2013) study are in line with the results of this study. They found that the peer review process is highly considered to be critical to scientific

research; the authors declared that the peer review process enhanced the last paper they published. The findings of other studies like Atjonen (2018), Kumar, Rafiq, and Imam (2011), Schwartz and Zamboanga (2009), Bunner and Larson (2012) also support the findings of this study.

In addition, from an academic writing instruction point of view, the findings of the current study give credit to what Paltrdige (2017) maintains with respect to the process of teaching and learning how to do peer review. He states that 'ways of doing' peer review depends on the specific context and discipline where the reviewers write reviews. In other words, both the values and expectations of the particular discipline must be met for "what 'counts' as research, how it should be framed, theorized, investigated as well as how it should be reported on" (p. 185). This contextual and particularist approach to academic writing will pay the price as the young and novice researchers move on the right track toward the right target (Hyland, 2015).

5. CONCLUSION

Two main implications could be implied from this study: one for journal authors, another for the reviewers. To produce a qualified academic work, the authors should be responsible for their work. Furthermore, the criteria obtained from our study besides those mentioned in other pertinent studies should be respected and followed. The authors should be trenchant in operating the prerequisites mentioned above of qualified academic production and the reviewers should be totally honest about their real opinion about the article under review, they should possess those characteristics mentioned in the review section (e.g., being faithful, timely, unbiased, responsible, knowledgeable, etc.).

In doing so, the reviewers must possess the necessary competence and mastery of their respective area(s) of interest and specialty, be the expert members of their scientific community, pay attention to the standards set by the particular journal, and consider all aspects of an acceptable and appropriate academic work to be publishable. In line with the agreed-upon sets of beliefs, values, and views of a specific discipline, what the reviewers write invoke certain structures of knowledge easily discernable by the members of that particular community (Frow, 2015). Therefore, the reviewers must take into account the disciplinary expectations of their particular field in the review reports that they prepare as the final outcome of their peer review practice.

Despite its pedagogical implications, this study is carried out on the reviewers from science and research journals. Thus, future studies should include other types of journals to enhance the study results. Our respondents were only Iranian reviewers. It is advisable to embrace international reviewers to examine their mindset towards peer-reviewing process.

References

Atjonen, P. (2018). Ethics in peer review of academic journal articles as perceived by authors in the educational sciences. *Journal of Academic Ethics, 6*(3). https://doi.org/10.1007/s10805-018-9308-3

Boggs, S. (2009). Paper, peer review, and vested interests [Guest editorial]. *IEEE Electrical Insulation Magazine, 25*(6). https://doi.org/10.1109/MEI.2009.5313703

Bornmann, L., & Mungra, P. (2011). Improving peer review in scholarly journals. *European Science Editing, 37*(2), 41-43.

Bunner, C., & Larson, L. E. (2012). Assessing the quality of the peer review process: Author and editorial board member perspectives. *American Journal of Infection Control, 40*(8). https://doi.org/10.1016/j.ajic.2012.05.012.

Bush, T. (2016). Understanding the peer-review process: Reject, revise, resubmit. In C. Sugrue & S. Mertkan (Eds.), *Publishing in the academic world: Passion, purpose and possible futures* (pp. 90–99). Routledge.

Chowdhry, A. (2015). Gatekeepers of the academic world: A recipe for good peer review. *Advances in Medical Education and Practice, 6*, 329-330. https://doi.org/10.2147/AMEP.S83887.

Committee on Publication Ethics (COPE). (2014). Retrieved from http://publicationethics.org/

Frow, J. (2015). *Genre* (2nd ed.). Routledge.

Guthrie, J., Parker, L. D., & Dumay, J. (2015). Academic performance, publishing and peer review: Peering into the twilight zone. *Accounting, Auditing, and Accountability Journal, 28*(1). https://doi.org/10.1108/AAAJ-11-2014-1871

Hadi, M. A. (2016). Fake peer-review in research publication: Revisiting research purpose and academic integrity. *International Journal of Pharmacy Practice, 24*(5), 309–310. https://doi.org/10.1111/ijpp.12307

Hames, I. (2012). Peer review in a rapidly changing landscape. In R. Campbell, E. Pentz, & I. Borthwick (Eds.), *Academic and professional publishing* (pp. 15–52). Chandos Publishing.

Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & Judson, K. C. (2010). Final report: Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines. Retrieved from http://escholarship.org/uc/item/15x7385g

Herbert, W., Marsh, & Ball, S. (1989). The peer review process used to evaluate manuscripts submitted to academic journals. *The Journal of Experimental Education, 57*(2), 151-169. https://doi.org/10.1080/00220973.1989.10806503.

Human Resource Development Quarterly. (2013). Published online in Wiley Online Library (wileyonlinelibrary.com), *Wiley Periodicals, Inc, 24*(4). https://doi.org/10.1002/hrdq.21176

Hyland, K. (2015). *Academic publishing: Issues in the challenges in the construction of knowledge*. Oxford University Press.

Johnson, B., & Christensen, L. (2019). *Educational research: Quantitative, qualitative, and mixed approaches* (7th ed.). Sage Publications.

Kumar, P., Rafiq, I., & Imam, B. (2011). Negotiation on the assessment of research articles with academic reviewers: Application of peer-review approach of teaching. *Higher Education, 62*(9), 315–332. https://doi.org/10.1007/s10734-010-9390-y

Leopold, S. S. (2014). Editorial: Peer review and the editorial process – A look behind the curtain. *Clinical Orthopaedics and Related Research, 473*(1-3). https://doi.org/10.1007/s11999-014-4031-x

Mandal, J., Giri, S., & Parija, S. C. (2012). Ethics of editorial and peer review. *Trop Parasitol, 2*(1), 4-5. https://doi.org/10.4103/2229-5070.97231

Mulligan, A., Hall, L., & Raphael, E. (2013). Peer review in a changing world: An international study measuring the attitudes of researchers. *Journal of the American Society for Information Science and Technology, 64*(8), 132–161.

Nature. (2006). Overview: Nature's peer review trial. Retrieved July 19, 2015, from http://www.nature.com/nature/peerreview/debate/nature05535

Nygaard, L. P. (2015). Publishing and perishing: An academic literacies framework for investigating research productivity. *Studies in Higher Education*. https://doi.org/10.1080/03075079.2015.1058351

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Sage Publications, Inc.

Paltridge, B. (2017). *The discourse of peer review: Reviewing submissions to academic journals*. Palgrave Macmillan UK.

Qing, F., Lifang, X., & Xiaochuan, L. (2008). Peer-review practice and research for academic journals in China. *Journal of Scholarly Publishing, 39*(4). https://doi.org/10.3138/jsp.39.4.417.

Trevino, L. K. (2008). Editor's comments: Why review? Because reviewing is a professional responsibility. *Academy of Management Review, 33*(1), 8–10.

Sarker, S. (2015). Publishing in leading journals: An overview for aspirant authors early in their career. In J. Liebowitz (Ed.), *A guide to publishing for academics: Inside the publish or perish phenomenon* (pp. 191–202). CRC Press.

Samraj, B. (2016b). Research articles. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 403–415). Routledge.

Schuhmann, R. (2008). Editorial: Peer review per physical review. *Physical Review Letters, 100*(5). https://doi.org/10.1103/PhysRevLett.100.050001

Schwartz, S. J., & Zamboanga, B. L. (2009). The peer-review and editorial system: Ways to fix something that might be broken. *Perspectives on Psychological Science, 4*(1). https://doi.org/10.1111/j.1745-6924.2009.01106.x.

Sciortino, J. E., & Siemens, D. R. (2013). The editorial process: Peer review. *Canadian Urological Association Journal, 7*(7-8). https://doi.org/10.5489/cuaj.1589

Seth, S., & Leopold, M. D. (2014). Editorial peer review and the editorial process – A look behind the curtain. *Clinical Orthopaedics and Related Research, 473*(1). https://doi.org/10.1007/s11999-014-4031-x

Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine, 99*(8), 178–182.

Spier, R. (2002). The history of the peer review process. *Trends in Biotechnology, 20*(7), 357–358.

Sposato, L. A., Ovbiagele, B., Johnston, S. C., Fisher, M., & Saposnik, G. (2014). Peek behind the curtain: Peer review and editorial decision making at stroke. *Annals of Neurology, 76*(2), 151–158. https://doi.org/10.1002/ana.24218.

Swales, J. M. (1996). Occluded genres in the academy: The case of the submission letter. In E. Ventola & A. Mauranen (Eds.), *Academic writing: Intercultural and textual issues* (pp. 45–58). John Benjamins.

Tan, Z. Y., Cai, N., Zhou, N., & Zhang, S. (2019). On performance of peer review for academic journals: Analysis based on distributed parallel system. *IEEE Access, 7*. https://doi.org/10.1109/2896978

Tite, L., & Schroter, S. (2007). Why do peer reviewers decline to review? A survey. *Journal of Epidemiology and Community Health, 61*(1), 9–12. https://doi.org/10.1136/jech.2006.049817

Appendix I: Open-Ended Questionnaire Items

How many reviews do you perform per year?
How many editorial boards do you serve on?
For how many years have you been writing reviews?
How much time do you spend on writing reviews?
What do you find most challenging about writing a review?
What do you find most straightforward about writing a review?
How did you learn to write a manuscript review?
On what grounds do you reject a manuscript?
On what grounds do you accept a manuscript?
In your view, what are the qualities of a good research paper?
When do you refer a manuscript to another reviewer?

Language Learning Opportunities in the Online Wild

Pouromid, Sajjad[1],* & Hosseininasab, Khatereh[2]

[1] Kansai University, Osaka, Japan
[2] Ph.D. Candidate, Osaka University, Osaka, Japan

Abstract

Recent research has heralded the role of social interaction in learning a second language. While earlier cognitive approaches to language learning attracted attention to individual factors involved in the process, social approaches regard learning as a pluralistic attempt which is materialized through participation. This shift in focus is important because it entails the study of language learning as it occurs in its natural habitat of social interaction rather than limiting it to formal educational settings. Mainstream SLA research has suffered from this limitation, with most studies in the field opting for experimental or quasi-experimental designs. Although informative in many respects, such studies lack the ecological validity to explore how learners approach the task of language learning in the real world. To address this issue, the study of language learning in the wild (outside formal educational settings) has gained momentum. The present study takes a similar approach to explore the affordances online learner-learner interactions may offer for language learning. Rather than tracking and measuring learning, it seeks to understand the potentials such interactions may have for language learning particularly because they happen in the absence of teachers. It builds upon data collected from video calls among Japanese and Taiwanese learners of English, transcribed and analyzed with a conversation analytic lens. The findings indicate that online interactions outside classroom provide learners with opportunities for extended negotiations for meaning, besides being a space for developing awareness for how interactions are structured in conversations taking place in the real world.

1. INTRODUCTION

Social interaction is often regarded as sine qua non of both first and second language learning. SLA research has embraced this understanding either by regarding social interaction as a space for providing learners with rich input and giving them the feedback that facilitates their cognitive learning process, or as "the site where learning as a socio-cognitive endeavor is collectively shaped through socially coordinated courses of activities (Eskildsen et al., 2019, p. 2). Contrary to this

* **Corresponding Author:** Pouromid Sajjad, Division of Promotion of Educational Development, Kansai University, Japan. **Email**: pouromid@kansai-u.ac.jp

proposition, however, most empirical SLA studies build upon data collected from experimental or quasi-experimental designs. Although informative in many respects, such an approach raises the question of ecological validity for mainstream SLA research: if social interaction is the primary source of language learning, how come our current understanding of how languages are learned by and large emanates from research taking place in its absence? It can be of course argued that there is a growing interest in research conducted in the language classroom setting, which indeed features more ecological validity than the laboratories where experimental research is carried out. It is important to take into consideration, however, that the interaction that takes place in the language classroom is also often highly structured, if not controlled for the very sake of research. Contrary to the classroom setting, the world of second language speakers defies any top-down structuring as it is, by definition, multilingual, multimodal, and a space where a multitude of semiotic resources coexist. As a response to this drawback, investigating language learning outside the traditional classroom setting and *in the wild* has taken momentum in recent years.

The metaphor of learning *in the wild* foregrounds the belief that cognition is socially situated, and hence its nuanced complexities can only be appreciated in learners' real-world interactions. In other words, as Hutchins (1995) puts it, the idea of cognition in the wild delineates studying it in its "natural habitat," which is the "naturally occurring culturally constituted human activity" (p. xiii). An obvious implication of this perspective for SLA research is the need for drawing on data collected from naturally-occurring language learner interactions outside the conventional language classroom (Firth & Wagner, 1997). Such a socially situated approach to SLA research in turn enables us to apprehend language learning in an ecologically valid manner, and understands, for instance, the affordances that naturally-occurring social interactions can offer and the subtle ways language learners utilize such affordances to transform mundane social encounters into learning environments (Kasper & Burch, 2016).

Exploring learning in its natural habitat of social interaction attracted attention after what Block (2003) calls "the social turn in SLA", which as the term suggests, highlights the social dimensions of learning. Contrary to earlier SLA studies, research after the social turn started to utilize audio and video recordings of learner interactions in the real world, often analyzing them with a conversation analytic (CA) lens among other approaches (Hellermann, 2008; Pekarek-Doehler, 2018; Wagner, 2015). The common principle that underpins most of such studies is that learning is embedded in the activities people jointly conduct in collaboration with others to assign meaning to the social world. In the case of SLA, this very principle shifts the object of learning from mastering the formal aspects of the linguistic system to developing and mobilizing semiotic resources to achieve and maintain intersubjectivity or mutual understanding in their everyday interactions. This latter object of learning also entails a redefinition of the concept of competence. If learning is socially situated and is achieved through carrying out various social actions, then competence cannot be regarded as a cognitive and intra- psychological ability. It is rather a matter of how multiple participants in a social action deploy semiotic resources in a contextually appropriate manner (Hellermann, 2011; Pekarek-Doehler & Pochon-Berger, 2018).

Against this backdrop, the present paper seeks to understand what affordances the *online wild* can offer language learners and how they may be able to jointly utilize such affordances to realize their language learning goals. While prior research on language learning in the wild has explored language learner interactions outside the classroom in physical environments such as study abroad and homestay programs (Dings, 2014; McMeekin, 2017; White, 2019), the present study draws upon data collected from language learners' online interactions outside classroom setting. This type of data was also used in previous research (Taleghani-Nikazm, 2019) with the goal of tracking

the development of interactional competence in L2 speakers, but as there has been a surge in the quantity of online interactions taking place around the globe since the current COVID-19 pandemic broke out early in 2020, there seems to be both the need and the opportunity to further explore this type of interaction among language learners.

2. THE STUDY

This study draws upon data collected from online video interactions of Japanese and Taiwanese learners of English as a foreign language. This online exchange was part of an online collaborative program between a Japanese and a Taiwanese university during which students were given various task to carry out in mixed nationality groups with the goal of improving their English proficiency. While some of these activities were done in the presence of teachers and can be counted as online classroom assignments, other activities were done solely by students and in the absence of the two teachers. In the latter format, the learner-learner interactions which comprise the data analyzed in the present paper, different groups of learners (usually 4-6 learners from both sides in each group) were given topics and a few prompts by their teachers and were required to meet up with their peers within the same group using a video-call application of their choice outside the formal class time. Although the topics of discussions were initially selected by the two teachers, members of each group could freely change the direction of their interactions with their own discretion. They were not instructed on issues such as how long their interactions were supposed to be or when and how they were supposed to complete the task. One student in each group, however, was asked to set up the online meeting, record it, and then share it with the two teachers. Students would not receive any feedback on the formal aspects of their language use during these virtual exchanges from their teachers, as the primary purpose of the task was to provide learners with an opportunity to practice using English in the real world and in the absence of the controlled interactional structure of the classroom.

Once the virtual exchanges were concluded and their video files were shared with the teachers, the researchers, one of whom was teaching the Japanese class, transcribed the data and used a CA framework to analyze them. It is also worth mentioning that these virtual exchanges were not planned for research purposes and would take place with or without the researchers' further analysis of the data resulting from them. This in turn means that the data used in this study would qualify as naturally-occurring, which is a requirement for doing CA research. Throughout the transcribed data, all participants will remain anonymous. Taiwanese learners will be referred to as TLs (TL1, TL2, etc.) and Japanese learners as JLs (JL1, JL2, etc.).

3. FINDINGS

In keeping with the aim of this study, instances of the affordances for language learning existent in the online learner- learner interactions outside the formal classroom setting were identified. In what follows, a number of such instances will be presented and accompanied by extracts from the transcribed interactions.

Extended Negotiations for Meaning

Negotiation for meaning is often regarded as essential for making interactions in L2 comprehensible. The importance of comprehensible input in language learning has been underlined in the language teaching field particularly in the input hypothesis (Krashen, 1985) and interaction hypothesis (Long, 1996). The process through which incomprehensible input becomes accessible to speakers of an L2 is generally referred to as negotiation for meaning. When L2 learners come across a communicative breakdown as a result of gaps in their L2 system, they need to devise compensatory strategies to overcome that breakdown. What happens meanwhile is in

fact a negotiation for meaning. This shows the value of such instances in the learning process since as Walsh (2014) posits in educational settings learning and teaching are materialized through interaction, and some even believe that not only interaction leads to learning but that it is learning. The analysis of the data in this study indicated that several instances of negotiations for meaning took place in the learner-learner interactions. Excerpt 1 illustrates some of such instances.

In the following sequence of interaction JL11 is interacting with five of her Taiwanese peers namely TL3, TL5, TL9, TL11 and TL8. Their discussion revolves around a photo which featured a notice at the entrance of a Japanese restaurant asking foreigners not to enter the place. A communicative breakdown occurs at the beginning of the sequence and it takes the participants quite a few turns to negotiate and find a way to solve that problem.

Excerpt 1: Negotiation for meaning in learner-learner interactions

JL11: So: I think the <u>fo</u>reigner want to eat Japanese food in Japanese restaurant (2.0)
TL3: huh? ((smiling)) (1.0)
JL11: so, the (.) limit[like]
TL3: uhuh]
TL5: hum=
TL8: =hum
JL11: like this is very: (.) I think is very (.)□bad ((pronounced as /bʌt/ instead of/bæd/))
6   **TL3**: bu-? =
TL5: =bat?
JL11: □bad ((pronounced as/bʌt/))(2.0)
TL3: □bat. ((looksuncertain))
TL8: BAT=
TL5: =bat (.) [oh
JL11: bad] (1.0)
TL3: [()
JL11: like] this □photo,
TL3: huh=
JL11: [is
TL5: ()]
JL11: is bad. (1.0)
TL3: o:h
JL11: yes.
TL5: e: (1.0)
TL9: b:at ((smiling))
TL3: ((laughs briefly)) e:
TL5: e: :(.) uhm (.) have another (.)word abo:ut bat?
TL11: bat=
JL11: e:h (.) [so:
TL9: do you] spell it? (3.0)
TL3: [can you
JL11: what?]
TL3: spell it? (1.0) spell (1.0)
TL8: can you spell it?
TL3: [hum

JL11: B] (.) A (.) D? (.) B   (.) A (.) [D
TL8: ye-] oh □ba:d oh □ba: d {o:h
TL3: Oh bad
JL11: >yeah yeah yeah yeah<

At the beginning of the sequence, JL11 states her disapproval of the content of the photo which is the subject of the discussion. After a few interjections by the Taiwanese learners that can be interpreted as their displays of listenership, in turn 5 JL11 says that asking foreigners not to enter a restaurant is bad. This soon turns out to be a source of communication difficulty. One may not expect a word as simple as "bad" to be the cause of a communicative breakdown, but apparently JL11's mispronunciation of the word leads to misunderstandings. This in turns leads to extended negotiations for meaning during which some of the participants test their hypotheses until an agreement is reached.

Towards the beginning of the negotiations TLs tend to hold themselves responsible for not understanding JL11's utterance and this can be seen in the way they keep repeating the word or other similar words or non-words (bu- or bat instead of bad for instance) throughout turns 6 to 12, while JL11 keeps repeating the original word. Since no agreement is achieved, however, JL11 gives up repeating and tries referring back to the photo in the assignment in turn 14. Following this, TL3 signals a change of state in his understanding by uttering the short token "oh" with a prolonged vowel. JL11 takes this as a sign of understanding and responds with a positive assessment in turn 19. Soon, however, in turn 23 and after TL3's brief laughter, she finds out that mutual understanding is not achieved yet. In turn 24, TL5 employs another strategy and utters a clarification request asking JL11 to offer an alternative word. Before JL11 finds a chance to respond to this request, however, TL9 comes up with a different strategy and asks for clarification through spelling out the source of trouble. TL9's question "do you spell it" is then repeated by TL3 which provides further evidence that she had not understood the word earlier in turn 19 despite uttering a state changer "oh." In turn 31, TL8 repairs TL9's question saying "can you spell it?" instead. Finally in turn 33, JL11 spells the word and this seems to resolve the situation since both TL8 and TL3 utter state changer tokens along with correctly pronouncing the trouble source which is confirmed by JL11 at the end of the sequence. Notice how these opportunities for testing different hypotheses and trying different tools for achieving understanding could be wasted if a teacher with higher interactional authority had repaired the source of trouble quickly after its utterance in turn 5.

Managing Closing Sequences

The beginning turns in an interaction are called opening sequences and the ending turns are referred to as closing sequences. Both of these sequences are important from different perspectives. However, the latter of the two seems to offer even more complexities. The reason is that while in an opening sequence all participants know that the interaction will any second start and expect its commencement, in a closing sequence there is often no clear hint as when the interaction is coming to a close. Furthermore, closing an interaction in the absence of a shared understanding regarding the appropriate time for a closure can be sanctionable. As a result, the speaker who intends to bring the conversation to an end often does so after a prelude to closure which is here referred to as a pre-expansion sequence. McLaughlin (1984) argued that there are usually three functions in closing a conversation, namely signaling that there is a movement towards a state of decreased access, expressing appreciation for the encounter and a desire for future contact, and summarizing what the encounter has accomplished. In the case of the data analyzed in the present study what

seemed to happen after a pre- expansion sequence was mainly an expression of appreciation and the desire for future contact. The following excerpt provides an example for how pre- expansions work to prepare the other speakers for a closure.

Excerpt 2: Pre-expansion in closing sequences

1	**TL5:** ehm (.) so: (.) is time for you to go to bed?
2	**TLs:** ((loud laughter))
3	**JL11:** ((looks surprised)) (3.0) [ehm
4	**TLs:** ((incomprehensible talk in Chinese among TLs))
5	**TL11**: ehm ((incomprehensible talk in Chinese))
6	**TL3**: ehm actually e:h we ha:ve som:e time limit (.) ehm in our school dorm
7	**JL11**: ((nods))
8	**TL3:** ehm we need to take bath an:d wash our clothes=
9	**TLs:** = ((brief laughter)) ((talk in Chinese)) =
10	**TL3:** twelve o'clock so maybe we need to:
11	**TL11:** we can chat next time
12	**TL3:** we (.) we can chat next time
13	**JL11:** oh ok ((laughter))
14	**TL5:** so: (2.0) see you next tim: e (.) by: e
15	**JL11:** [by:e
16	**TLs:** by: e]

In this excerpt, JL11 and a couple of her Taiwanese peers are discussing local specialties in their hometown. After completing the assignment, TL5 asks a question which does not fit into the content of its preceding turns. As the rest of the interaction unfolds, however, this question can be better understood.

As can be seen in the excerpt, TL5 asks JL11 whether she has to go to bed in turn 1. This is followed by the other TLs' laughter implying that they may know why TL5 is asking this question. JL11 looks surprised and is seemingly not sure whether she has understood the question well. TL3's comments in turns 6, 8 and 10, however, reveal the real purpose of the question asked by TL5 in turn 1. TLs seem to have a time limit in their dormitory and have to end the video call soon. Yet, since they perceive ending the call without prior preparations as a dispreferred action, they initiate a pre-expansion sequence before the actual closing sequence. The plan would have worked better if JL11 had given a positive response to the question saying that she had to go to bed, but since she did not say so, TL3 had to explain why the question was asked. The closing sequence is then initiated by TL11 who expresses the TLs' will to continue the chat next time. The combination of the pre-expansion sequence and the closing sequence brings about a smooth closure to the sequence.

Turn Distribution Bias

Turn taking and turn distribution are two key areas in the study of social interactions. In the SLA context, they may also be an indicator of learners' interactional competence. Taking turns by learners in an educational setting also allows them to have autonomy in the learning process. The opposite can also stand true. That is, a rigid turn distribution pattern by a teacher in a classroom may well limit the learners' chance to exercise autonomy by self-selecting for upcoming turns. In the present study, however, teachers were absent in the data and this could potentially mean that learners had the chance to take turns more freely. It was not necessarily so, nevertheless. In fact,

the analysis of the data showed that while in some instances self-selection took place and turn taking was done in a more fluid manner, in many others there was a bias in the distribution of turns. As shall be seen in the two following excerpts, one of the learners would take on a teacher role in the interaction thereby disrupting voluntary turn taking by distributing turns among the other participants. The teacher-figure was often the one who was in charge of starting the video call and inviting the other members. What is even more important is the way other participants orient to the position the learner/teacher-figure assumes for him/herself by treating him/her as a participant with teacher responsibilities.

The following sequence of interaction happens in the absence of Japanese participants. Five Taiwanese learners join a video call hosted by TL18 to do an assignment. The topic of the discussion was assigned by one of the teachers and dealt with the issue of foreigner-friendly restaurants in Japan and Taiwan. As in this particular case, Japanese learners are not present, all discussions are about the Taiwanese context.

Excerpt 3. Turn distribution by a teacher figure in L/L interactions

1 **TL18:** so: ((coughs)) hello everyone ((laughs))
2 **TL7:** hello: [((waves at the camera))
3 **TL14:** ((Waves at the
4 **TL17:** camera)) hi:] ((Waves at the
5 **TL18:** camera)) ((waves back)) I'm
6 **TL17:** glad to call the roll (.) so: hum=
7 **TL18:** =TL17 ((TL17's
8 **TL17:** name)) hey yeah ((raises
9 **TL18:** his hand)) u:h TL2 ((TL2's
10 **TL2:** name)) ((raises her hand while smiling))
11 **TL18:** a: nd u:h ((looks away from the camera as if trying to remember something)) and who? TL7 ((TL7's name))
12 **TL7:** [((raises her hand but revokes the action halfway as TL14 takes up the next turn))
13 **TL14:** ↑TL14 ((her own name, notably louder))
14 **TL18:** uh ((laughs))
15 **TL17:** [((laughs))
16 **TL7:** ((laughs))]
17 **TL18:** hey TL14 ((her name)) yes of course TL14 ((her name))
18 **TL14:** ((raises her hand)) (yeah)
19 **TL18:** ok so=
20 **TL7:** =(what) about me? =
21 **TL18:** =let's get (cracking)=
22 **TL7:** what about me?
23 **TL14:** yeah ((pointing to TL7))
24 **TL18:** yeah, I ↑said TL7 ((her name))
25 **TL7:** oh, uh ok ((raises her hand and waves at the camera))
26 **TL18:** ok so (.) first question ((brings his head closer to his monitor to read the question from his screen)) (2.0) are restaurants oh fu- (.)are [restaurants
27 **TL2:** what?]
28 **TL18:** in Japan and Taiwan foreigner friendly >give plenty of examples to support your stance<. hhh (3.0)

29 **TL14:** ((raises her hand and waves))
30 **TL18:** ((notices TL14's gesture, raises his hand and waves back while smiling))
31 **TL17:** hey TL14 ((her name))
32 **TL14:** oh (.) oh me ((looks at her notes briefly and then looks back at the camera)) ↑yes (.) as my sister's experience in Chinese restaurant ((looks at her notes from time to time)) they will provide many kinds of tableware such as knife and fork for the foreigners not only chopsticks (1.0) is is kind of is kind to foreigner ((thumbs up))
33 **TL18:** ((makes a funny gesture putting her fingers around her eyes and rolling her eyes))
34 **TL7:** ((laughs while pointing to her screen))
35 **TL14:** ((laughs))
36 **TL18:** ((shows thumbs up with both hands while laughing))
37 **TL17:** yeah
38 **TL7:** ((pointing to herself)) my turn my turn
39 **TL18:** ok TL7 ((her name)) your turn. hhh
40 **TL7:** uh I think [that
41 **TL17:** ((laughs))]
42 **TL7:** there's a restaurant called (Ting Tai Fong) [a:nd is
43 **TL18:** Oh, I hate it]
44 **TL7:** foreign friendly is ↑foreign friendly for foreigners because the menu has some English and Japanese (1.0) [so foreigners
45 **TL18:** and it's expensive]
46 **TL7:** come here to if they want to eat there, they could see the menu easily
47 **TL18:** expensive (1.0) expensive
49 **TL7:** good ((thumbs up))
50 **TL18:** expensive [expensive
51 **TL7:** delicious]
52 **TL18:** yes, delicious and expensive
53 **TL7:** delicious (shoronpo) ((Taiwanese food))
54 **TL18:** ((laughs)) (shoron) best thing dumpling
55 **TL7:** ((laughs))
56 **TL18:** I think so ok so=
57 **TL7:** =ok
58 **TL18:** next question (7.0) ((looking for the question on his screen while bringing his head very close to it and the camera))
59 **TL14:** ((laughs))
60 **TL18:** (reads from his screen)) what can be done to make restaurants friendlier to foreigner visitors?
61 **TL2:** ((raises her hand)) ↑me
62 **TL18:** o:h a: nd e:hm (1.0)
63 **TL17:** TL2 ((TL2's name)) =
64 **TL18:** =TL2 ((her name)) ye: s TL2 ((her name)) (1.0)
65 **TL2:** u:h I think we can cha- change the staff (.) u:hm make them (.) learn (.) some (.) foreign (.) language (.)and give some picture on the menu=
66 **TL18:** O:h pictures
67 **TL2:** yeah
68 **TL18:** good advice (2.0) >bravo< ((laughs))

The sequence begins with TL18's greeting the other participants as the host of the video call. Soon in the sequence and in turn 5, TL18 explicitly positions himself as the facilitator of the interaction by announcing that he is "glad to call the roll." In an educational setting, this task is often performed by a teacher and therefore TL18 creates this assumption that he will be playing the role of a teacher though all other participants can acknowledge that he is not really a teacher. TL18's announcement in turn 5 corresponds with what he does in the coming turns as well. He goes on by calling other participants' names and waits for them to react (turns 7 to 10, for instance). It is also interesting that other participants tacitly agree with the way TL18 has positioned himself. TL7, as a case in point, reminds TL18 in turn 20 that he has not called her name similar to what might happen in a real classroom in the presence of a teacher.

After calling everyone's names TL18 moves the interaction to the next phase in turn 26. Similar to what a teacher might do, he starts his turn with the short token "ok" to indicate a change of topic and then reads out the first question of the assignment. The question comes to an end in turn 28 after which a three-second pause emerges. This can be interactionally interpreted as a chance for self-selection by the next speakers in the sequence and this is exactly what happens in turn 29. TL14 volunteers to respond to the question. However, treating her turn as a response to the teacher-figure's turn, she raises her hand waving at the camera and waits until TL18 explicitly allows her to take the turn. In other words, although self-selection takes place at this point, it is mediated through permission seeking which can be a sign of perceived power imbalance in turn taking. Once again, this reminds one of the turn allocation patterns that takes place in a classroom and in the presence of a teacher. It is also worth mentioning that it is not TL18's observable outward behavior that positions him as a teacher. Notice, for example, how in turn 33 he makes funny gestures that make other participants laugh. This might be very unlikely for a teacher in an institutional setting. It is rather TL18's interactional moves realized through the way he distributes turns as well as the way other participants orient to his tacitly agreed upon right to do so that makes him look and sound like a teacher. Other examples for turn allocation can also be found in the excerpt. TL7 in turn 38 also nominates herself for the next turn but only starts to speak when TL18 explicitly invites her to do so. Turns 63 and 64 follow a similar pattern too. The sequence comes to an end with TL18's explicit positive feedback which provides even further evidence for the way he has positioned himself. Explicit positive feedbacks are uttered in the feedback of a tripartite IRF sequence by the same speaker who has initiated a given sequence and, in this case, distributed the turn to other participants.

All of the examples given above involve instances of mediated self selection by potential next speakers. That is to say, in all of those cases, the next speaker would self select at a transition-relevance place (TRP), but would not actually start her turn without seeking permission from the teacher figure. The following excerpt illustrates what might happen if the next speaker fails to seek permission before taking a turn while there seems to be a perceived imbalance in turn taking rights by the speakers. The excerpt takes place within the same video call as the previous one with the same participants.

After TL18 shows his intention to move to the next question of the assignment by uttering the transition marker "ok", TL7 briefly states what the next question is by referring to its number in the assignment rubric. After a short pause which introduces a TRP, TL17 self-selects and without waiting for TL18's indication of permission starts his utterance in turn 3. TL17's turn overruns with TL18's throughout turns 3 to 5. As TL18's reaction in turn 6 indicates, TL17's undue initiation of his turn is not tolerated by the teacher- figure who explicitly asks TL17 to "wait" with

a higher pitch. This is often referred to as a sanction that the speaker who made an undue attempt to take a turn has to endure. TL17's response to this also indicates his tacit agreement with TL18's superiority in allocating turns.

Excerpt 4. Sanctions in undue turn taking in L/L interactions

TL18: ok (.) move on (.)
let's move on (.) the next question
TL7: number four (2.0)
TL17: is my [question
TL18: the last] [question
TL17: is e:h] =
TL18: □wait
TL17: ((laughs)) ok

4. DISCUSSION

In what follows, the three issues presented above, namely extended negotiations for meaning, managing closing sequences, and turn distribution bias will be discussed in light of what is already known from the relevant literature.

Extended Negotiations for Meaning

The concept of negotiation for meaning has been around in applied linguistics for quite a few years. It fits well within the cognitive accounts of language learning and was first introduced in Long's Interaction Hypothesis (1996). Its theoretical tenets, however, can be traced back to mid-80s and Krashen's (1985) Input Hypothesis. For Krashen, the key to the acquisition of a second language was exposure to comprehensible input. Not all input is comprehensible for a language learner. For input to meet the criterion of comprehensibility it had to be just above a learner's current language level. Krashen called this level i+1. He believed that if the incoming information meets this criterion, it will be first comprehended and then acquired. Krashen's proposition was considered as an intuitive suggestion at the time, yet it suffered from a practical issue. Since i+1 would be different for each individual learner and also for one learner from one time to another, how could one make sure whether the input was comprehensible? In other words, how could a teacher identify the "i" level in order to generate the i+1 input? These were the questions Long addressed in his Interaction Hypothesis later on.

Long argued that the best way to achieve comprehensible input is through interactional adjustments. Interactional adjustments mean that language learners can make incomprehensible input comprehensible by negotiating meaning through which clarification and modifications are made in the information until mutual understanding or intersubjectivity is achieved. At this moment, Long argues, the information is brought into the learners' i+1 range. Individual learners with different language levels can modify the incoming information through negotiations until complete comprehension is achieved.

Krashen's Input Hypothesis and Long's Interaction Hypothesis served as theoretical underpinnings for later research. Although there have been slight differences in the definitions proposed for negotiation for meaning, there seems to be a shared understanding in the field about what generally constructs it. For instance, for Pica (1992, p. 200) negotiation for meaning was "an activity that occurs when a listener signals to the speaker that the speaker's message is not clear

and the speaker and the listener work linguistically to resolve this impasse." Gass and Selinker (1994, p. 209) also argued that it comprises "instances in conversation when participants need to interrupt the flow of the conversation in order for both parties to understand what the conversation is about." Smith (2005) on the other hand, stated that negotiation for meaning is an explicit indication of non-understanding and the subsequent attempts to resolve it. From an interactional perspective also van der Zwaard and Bannink (2014) argued that negotiations for meaning are a series of conversational turns that start due to an absence of understanding and continue until comprehension is achieved. Despite differences in their focus and terminology, these definitions have a lot in common: they all agree that negotiation for meaning is a response to a sort of what Long (1996) called a communicative trouble or what Mackey, Gass and McDonough (2000) called a communication breakdown which continues until the problem is resolved.

Research on negotiation for meaning has shown how it can lead to the uptake of new linguistic knowledge by the learners. A considerable number of studies in this area have investigated different task types that are more likely to generate negotiations and lead to linguistic uptake in the language classroom (Nakahama et al., 2001, for instance). A relatively smaller number of studies have also shed light on the interactional intricacies involved in negotiations for meaning (Zheng et al, 2009) studying turn taking and repair organization during negotiations both in traditional instructional settings and in online learner-learner interactions. Although more research has been conducted in instructional settings highlighting the role of teachers in managing negotiations for meaning, studies focusing on how learners learn from each other have gained popularity in more recent years as well. Among these studies, there are those that are closer in scope to the present paper particularly because they have studied learner-learner negotiations for meaning in technology mediated online settings. This latter group of studies has shown how various modes of online communication create different opportunities for negotiation for meaning (Yuksel & Inan, 2014). Studying online video interactions among L2 learners Sert and Balaman (2018), for instance, found that learners negotiate very often different aspects of the task including both linguistic and managerial aspects of them whenever there are problems of shared understanding generating regulations through repair initiation and accomplishment that assist them maintain understanding.

One of the findings of the present research was that extended negotiations for meaning including multiple lengthy instances of repair. Taking into account the theoretical principles reviewed in the above, it can be argued that deeper and lengthier instances of negotiations create better opportunities for language learners. Similar to what data-driven studies on negotiation for meaning counted out here have found, it was also found in this study that such instances initiate as a result of a gap in understanding or a communication problem. Such problems could take different syntactical, phonological, and lexical forms. As a result of lengthy and extended negotiations learners would generate hypotheses, test them and then either approve or revise them until intersubjectivity was achieved. Such cases often occur less frequently in teacher-led interactions and this can be supported with what is currently known about the interactional structure of language classes. Sert (2015), as a case in point, argues that teacher-initiated teacher repairs are very common in teacher-led interactions. Repairs as such correct learner mistakes on the spot and limit the space for negotiations among learners. In the absence of teachers, however, as was found in this study self-initiated repair may be more dominant resulting in lengthier and deeper negotiations for meaning.

Managing Closing Sequences

The way the endings of video calls were managed by the learners in learner- learner online interactions was another finding in the present research. In CA studies, conversation is regarded as a system, which is comprised of different parts. Studies on telephone conversations, for example, have found that there are opening and closing sequences in them (Wong & Warring, 2010). These sequences allow participants to start and end conversations as smoothly as possible. There are also normative orientations towards these sequences, meaning that the absence of an appropriate opening or a closing sequence might create communication problems and even sanctions for the speaker who failed to initiate such sequences. This point is closely related to preference organization in interactions as well. Inspired by the pioneering works of Pomerantz (1984) and later on Schegloff (2007), research on preference organization has shown how while participants evaluate certain responses to a first pair part (FPP) as preferred, other responses might be considered as dispreferred. As a case in point, accepting an invitation is generally considered as a preferred response, while declining one is dispreferred. That is why acceptance responses are given very quickly and in an unmarked manner while rejection is usually prefaced with justifications or pauses. As far as telephone conversations are concerned, closing the talk without prior indications of the fact that it is going to be closed soon is also considered as a dispreferred action. Avoiding this scenario, therefore, may indicate a speaker's level of contextual awareness and interactional competence.

Compared with research on telephone conversations between L2 learners, fewer studies have investigated video calls among second language speakers. Although there are structural similarities between these two types of telecommunications, there are differences as well. Apart from the obvious fact that in the latter case there are both audio and visual modalities, there may be more than two participants in video calls as well. Similarly, more research has been done on L1 telephone and video conversations in comparison with L2 conversations. What we already know about closing sequences in L1 telephone conversations, however, might guide us in analyzing closings in multiparty video calls, too.

In native speaker interactions, for instance, it is known that closings in telephone conversations come after pre- closing sequences a dominant feature of which is exchanges of short utterances such as okay, alright, good or the like. These utterances often appear after inter- turn pauses and when propositional meanings of the previous turns are already understood. Take this example from Schegloff and Sacks (1973):

A: O.K.
B: O.K.
A: Bye bye.
B: Bye.

The argument here is that to mitigate the possible dispreferred bearing of an abrupt closing, speaker A provides hint for the upcoming action of closing. Once speaker A receives speaker B's approval in the form of repeating his/her repetition of the short utterance, speaker A initiates the closing sequence. Notice that speaker B could have potentially opted not to allow A to end the conversation by saying "by the way" for instance. The closing here is therefore constructed by both speakers. This may seem very straightforward and native speakers of any language may take it for granted. In a second language, however, no matter how simple it may seem, it can be challenging. As Wong and Warring (2010) posit, second language

learners "do not necessarily know how to get out of a conversation or how to extend it in a second language" (p.11). Not knowing how to do so, therefore, they may either end up sounding awkward or impolite or devising their own strategies to end a call smoothly. The latter was the case with the learner-learner interactions in the present research.

The analysis of the closing sequences in learner-learner interactions in this study showed that the participants(s) who wanted to end the call would use two strategies both happening before the actual closing section. First, they would initiate a pre-expansion sequence justifying their upcoming closure of the talk by providing reasons. Doing so, they could actually test the water and learn about the other participants' intentions regarding either closing or continuing the video call. Second, they would offer to have another video call in the future to pick up where they leave off in the current one. The strategy of offering to do something in the future in pre-closing sequences was also found in the study conducted by Curl (2006). Overall, both these strategies served the purpose of informing the other participants' that a closing sequence would be ahead. This, as explained above, could mitigate the dispreferred bearing of an abrupt closing.

Turn Distribution Bias

Turn taking plays a pivotal role in social interactions. The amount of interaction taking place in any given conversation is closely correlated with how many turns are taken by the participants in that conversation. In the case of the second language classroom, it becomes important in another way as well. As discussed earlier, a number of language learning theories emphasize that learning is materialized through participation. Without a turn taking, there will not be any participation. Taking turns is not always easy for language learners, however. To take a turn, particularly when one is not selected as a next speaker, requires possessing the linguistic and sociolinguistic knowledge of how an interaction works. For instance, one has to be able to anticipate when a TRP is going to emerge to be able to take a turn without sounding improper or awkward (Pekarek- Doehler & Pochon-Berger, 2015). And to make things worse, all this has to be done in milliseconds. Cognitive demand for taking turns is high for second language learners and that is why learners with lower proficiency levels are often reported to have difficulty managing turns at talk (Carroll, 2004). The literature on classroom interaction, however, has shown that teachers have traditionally played important roles in making it easier for language learners to take turns.

In the language classroom, turn taking follows a particular system in which it is the teacher who often selects the next speaker. The reason for such a tendency is at least twofold. First, given the fact that there is limited time in a classroom, there is a tendency among teachers and learners to be internationally economical (Kääntä, 2010). Classrooms follow syllabi that need to be covered and allowing all learners to take turns at any moment during the class time might not simply be feasible. The language classroom is an "institutional context in which participants come together to achieve the specific goal of teaching and learning" (Garton, 2012, p. 29) and therefore classroom discourse is a form of institutional talk following its own rules and regulations. Second, there are asymmetrical role relationships between teachers and learners. This is especially evident in the turn taking system: most often it is the teacher who has more interactional power and decides who speaks when (Walsh, 2006). This, of course, does not mean that learner self- selection and initiation do not take place.

Taking these two reasons regarding why teachers often select next speakers in the classroom setting, it could be therefore assumed that learner-learner interactions outside the context of the language classroom and in the absence of teachers would feature more self-

selection instances. The analysis of the data in this study, however, showed that this is not necessarily the case. Turn taking in learner-learner interactions was not always fluid and voluntary. In other words, instead of making attempts to take turns in their interactions, the participants in this study seemed to tacitly orient to one of their peers in the interaction as a teacher. It was this teacher-figure who distributed the turns among other participants then.

The decision regarding who had the responsibility to manage the turns was not explicitly made known, but both the teacher-figure and other participants seemed to accept this division of roles and orient to it accordingly.

What the data also showed was that the teacher-figure was often the member who would set up the video call and invite others to join. However, this has to be treated with caution since there may well be other factors not identifiable with the design of the present research. One hypothesis would be that the learners who set up video calls and subsequently took up the teacher-figure role were perceived to have higher English proficiency levels by their peers. As there is no information regarding the perceptions of the participants in this study, such ideas will remain hypotheses, however. What the data actually shows is that constraints of institutional talk on turn taking (Garton, 2012; Walsh, 2006) do not necessarily relax in the absence of teachers and when learner-learner interactions take place outside the physical classroom setting. Learners may co-construct the classroom context and hence follow the regulations of the classroom institutional talk outside the classroom as well. This is for sure in line with the pioneering works of Sacks, Schegloff, and Jefferson (1974) who argued that turn taking is "locally managed, party- administered, and interactionally controlled" (p. 727).

5. CONCLUSION

Building upon data collected from online learner-learner interactions among Japanese and Taiwanese learners of English as a foreign language, the present study explored the affordances of such out-of-classroom interactions for language learning. The findings indicated that in the absence of teachers, such interactions offer learners the opportunity to negotiate for meaning to resolve their interactional problems. The data also showed how through using language for real-world purposes and in meaning- focused and goal-oriented interactions, learners can become aware of the interactional structure required to produce contextually appropriate utterances. The absence of teachers from these interactions, however, did not lead to the co-construction of a more fluid turn-taking pattern. The fact that one of the participants would take on a teacher's role and distribute turns among other participants meant that there were few opportunities for participants to self- select for upcoming turns and initiate new sequences. This might demonstrate the need for raising awareness in L2 learners about how interactions outside the classroom, or interactions in the wild (online or otherwise) can be different from those taking place within the walls of a language classroom. Learners who have been accustomed to rigid turn-taking patterns of a classroom may find it difficult to venture other patterns even when the restrictions of the institutional settings no longer apply. Online platforms can offer a viable solution to this problem by providing space for L2 learners to be exposed to language use in the wild.

References

Block, D. (2003). *The social turn in second language acquisition*. Edinburgh University Press.

Carroll, D. (2004). Restarts in novice turn beginnings: Disfluencies or interactional achievements? In R. Gardner & J. Wagner (Eds.), *Second Language Conversations* (pp. 201–219). Continuum.

Curl, T. S. (2006). Offers of assistance: Constraints on syntactic design. *Journal of Pragmatics, 38*, 1257-1280. https://doi.org/10.1016/j.pragma.2005.09.006

Dings, A. (2014). Interactional competence and the development of alignment activity. *The Modern Language Journal, 98*(3), 742-756. https://doi.org/10.1111/modl.12119

Garton, S. (2012). Speaking out of turn? Taking the initiative in teacher-fronted classroom interaction. *Classroom Discourse, 3*(1), 29-45. https://doi.org/10.1080/19463014.2012.666021

Gass, S., & Selinker, L. (1994). *Second Language Acquisition: An introductory course*. Lawrence Erlbaum Associates.

Hellermann, J. (2008). *Social actions for classroom language learning*. Multilingual Matters.

Hellermann, J. (2011). Members' methods, members' competencies: Looking for evidence of language learning in longitudinal investigations of other-initiated repair. In J. K. Hall, J. Hellermann, & S. Pekarek-Doehler (Eds.), *L2 interactional competence and development* (pp. 147–172). Multilingual Matters.

Hellermann, J., Eskildsen, S. W., Pekarek-Doehler, S., & Piirainen-Marsh, A. (2019). Conversation analytic research on learning-in-action: The complex ecology of second language interaction "in the wild". Springer Nature. https://doi.org/10.1007/978-3-030-24799-8.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Kääntä, L. (2010). Teacher turn-allocation and repair practices in classroom interaction: A multisemiotic perspective (Unpublished doctoral dissertation). University of Jyvaskyla.

Kasper, G., & Burch, R. A. (2016). Orienting to focus on form in the wild. In R. A. van Compernolle & J. McGregor (Eds.), *Authenticity, language, and interaction in second language contexts* (pp. 198–232). Multilingual Matters.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In R. William & B. Tej (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*(4), 471-497. https://doi.org/10.1017/S0272263100004024

McLaughlin, M. (1984). *How talk is organized*. Sage.

McMeekin, A. L. (2017). L1-L2 speaker interaction in a study abroad setting: Communication strategies, word searches, and intersubjectivity. *SARSLAIA, 2*(2), 263-294.

Nakahama, Y., Tyler, A., & Van Lier, L. (2001). Negotiation of meaning in conversational and information gap activities: A comparative discourse analysis. *TESOL Quarterly, 35*(3), 377–405. https://doi.org/10.2307/3588036

Pekarek-Doehler, S. (2018). Elaborations on L2 interactional competence: The development of L2 grammar-for-interaction. *Classroom Discourse, 9*(1), 3–24. https://doi.org/10.1080/19463014.2018.1448227

Pekarek-Doehler, S., & Pochon-Berger, E. (2015). The development of L2 interactional competence: Evidence from turn-taking organization, sequence organization, repair organization, and preference organization. In T. Cadierno & S. Eskildsen (Eds.), *Usage-Based Perspectives on Second Language Learning* (pp. 233-270). De Gruyter Mouton.

Pekarek Doehler, S., & Pochon-Berger, E. (2018). L2 interactional competence as increased ability for context-sensitive conduct: A longitudinal study of story openings. *Applied Linguistics, 37*(4), 555–578. https://doi.org/10.1093/applin/amw021

Pica, T. (1992). The textual outcomes of native speaker-nonnative speaker negotiation: What do they reveal about second language learning? In C. Kramsch & S. McConnell-Ginet (Eds.), *Text and Context: Cross-Disciplinary Perspectives in Language Study* (pp. 198-213). D. C. Heath and Company.

Pomerantz, A. (1984). Pursuing a response. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 152-163). Cambridge University Press.

Sacks, H. E., Schegloff, E., & Jefferson, G. (1974). A simplest systematic for the organization of turn-taking for conversation. *Language Awareness, 50*, 696-735. https://doi.org/10.2307/412243

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in Conversation Analysis*. Cambridge University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 7*, 89–327.

Sert, O. (2015). *Social interaction and L2 classroom discourse*. Edinburgh University Press.

Sert, O., & Balaman, U. (2018). Orientations to negotiated language and task rules in online L2 interaction. *ReCALL, 30*(3), 355-374. https://doi.org/10.1017/S0958344017000341

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly, 39*(1), 33–58. https://doi.org/10.2307/3588451

Taleghani-Nikazm, C. (2019). Ohja, ja, ja: Providing the appropriate next relevant action in L2 interaction. In R. M. Salaberry & S. Kunitz (Eds.), *Teaching and testing L2 interactional competence* (pp. 125-141). Routledge.

van der Zwaard, R., & Bannink, A. (2016). Nonoccurrence of negotiation of meaning in task-based synchronous computer-mediated communication. *The Modern Language Journal, 100*(3), 625–640. https://doi.org/10.1111/modl.12350

Wagner, J. (2015). Designing for language learning in the wild: Creating social infrastructures for second language learning. In T. Cadierno & S. Eskildsen (Eds.), *Usage-based perspectives on second language learning* (pp. 75–101). De Gruyter.

Walsh, S. (2006). *Investigating classroom discourse.* Routledge.

Walsh, S. (2014). Developing classroom discourse competence. *Language Issues, 25*(1), 4-14.

White, K. (2019). Interactional competence and study abroad: Empirical methods, findings, and pedagogical implications. In R. M. Salaberry & S. Kunitz (Eds.), *Teaching and testing L2 interactional competence* (pp. 192-212). Routledge.

Wong, J., & Waring, H. Z. (2010). *Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers*. Routledge.

Yuksel, D., & Inan, B. (2014). The effects of communication mode on negotiation of meaning and its noticing. *ReCALL, 26*(3), 333–354. https://doi.org/10.1017/S0958344014000165

Zheng, D., Young, M. F., Wagner, M. M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *The Modern Language Journal, 93*(4), 489–511. https://doi.org/10.1111/j.1540-4781.2009.00927.x

Disciplinary and Cross-Cultural Variation of Stance and Engagement Markers in Soft and Hard Sciences Research Articles by Native English and Iranian Academic Writers: A Corpus-based Analysis

Haniye Seyri[1,*] & Saeed Rezaei[2]

[1] University of Tehran, Tehran, Iran
[2] Sharif University of Technology, Tehran, Iran

Abstract
Drawing on a corpus-based approach, this study analyzed two different sub-corpora including Non-Native English-Speaking (NNES) and Native English-Speaking (NES) sub-corpus. There were 60 research articles from soft sciences including Applied Linguistics, Sociology, Economics and hard sciences including Chemical Engineering, Electrical Engineering, and Biology. To examine the frequency of stance and engagement markers in the two sub-corpora separately, MAXQDA software was utilized. Several Chi-square tests were run to investigate the differences found in the frequencies of the two groups. The results demonstrated that writers of different fields of study and from different cultural backgrounds exerted varying degrees of authorship and interaction in their texts. Regarding disciplinary variation, it was found that the researchers in soft disciplines used more stance and engagement markers than the ones in hard disciplines. With regard to cross-cultural variation, native academic writers preferred to draw more on interactional markers than non-native Iranian academic writers. The findings of the present study offer implications to academic writers from different fields of study and different cultural backgrounds so that they become cognizant of their own presence in texts and their interaction with readers based on the use of stance and engagement markers. The results can also be implemented in EAP/ESP courses and syllabi.

1. INTRODUCTION

Academic writing has been viewed as a social process in which authors can exert authorship in their texts and build rapport with their readers. This enterprise has been considered as a persuasive endeavor which puts emphasis not only on constructing texts, but also on establishing social relations using language (Hyland, 2005a, p. 173). Accordingly, one of the features of scientific discourse in the current academic milieu is how authors represent themselves and interact with their readers, with the former achieved via stance markers and the latter by engagement markers (Hyland, 2005a, 2005b). By applying these linguistic devices, writers can make their presence

---

* **Corresponding Author:** Saeed Rezaei, Linguistics Languages and Linguistics Center Sharif University of Technology. **Email**: srezaei@sharif.edu

more or less tangible in their texts and establish effective interaction with their audience depending on a myriad of factors including the rules of the disciplinary community to which they belong (e.g., Hyland, 2005a; McGrath & Kuteeva, 2012) and the L1 and cultural context in which they write (e.g., Çandarlı, Bayyurt, & Marti, 2015; Xu & Nesi, 2019; Yang, 2014). According to Hyland (2008a), the use of stance and engagement markers is a context-dependent matter in a way that authors draw on these markers based on the context, be it discipline or culture, in which they are positioned.

The claim that discipline exerts influence on the linguistic choices of academic writers has been corroborated by many studies (Hyland, 2001, 2002a, 2005a). According to Hyland (2002b), ―academic writing is not a single undifferentiated mass, but a variety of subject- specific literacies‖ (p. 352). In this sense, academicians draw on various linguistic devices that are compatible with the expectations of their disciplinary communities (Işık-Taş, 2018). Furthermore, the realization of science-based knowledge is profoundly rooted in culture- specific beliefs and norms resulting in a variety of intellectual styles of presentations and interactions in scientific discourse (Abdollahzadeh, 2011; Shaw, 2003). In this specific context and following the literature (e.g., Dontcheva-Navratilova, 2021; Shaw, 2003; Xu & Nesi, 2019), L1 is considered as the representation of culture since L1 writing strategies, L1 rhetorical structures, and cultural conceptualizations are evident in L2 written discourse (Connor, 1996, 2004; Dahl, 2004; Kaplan, 1966; Sharifian, 2009). In this regard, L1 plays a pivotal role in the way writers delineate an expression of themselves as members of a disciplinary community and in the way they demonstrate interaction with their audiences. Along the same line, a number of studies have emphasized that the linguistic choices of writers are highly influenced by their L1 and this claim can be supported by studies that compare authors writing in English as an additional language and the ones writing in English as their first language (Dontcheva- Navratilova, 2021; Lafuente-Millán, 2014; Xu &Nesi, 2019).

Despite the plethora of studies on stance and engagement markers, little is known about the use of stance and engagement markers in English-medium research articles (henceforth RAs) of different disciplines written by Iranian NNES academics and also the differences between English native speakers and Iranian non-native speakers of English in terms of the use of these markers at the same time. Considering the latter, it is indicated that the variation in rhetorical structures of texts written by non-native speakers of English is not merely discussed in terms of grammatical and semantic features of different languages rather it can be attributed to differences in reader-writer interactions created by culture-specific conventions (Hyland, 2008b) and the ―phenomenological differences between the cultures‖ in which authors are engaged (Kaplan 1976. p. 17). Accordingly, an in-depth analysis of how L1 and cultural background of writers exert influence on the construction of L2 texts merits attention (Atkinson, 2003; Connor, 2004). The current study will contribute to this line of enquiry by adopting a corpus-based approach. Therefore, the aims of this study are twofold: 1) To investigate the use of stance and engagement markers in Iranian-authored English RAs across six different fields of study, 2) To examine the influence of culture on the use of stance and engagement markers by considering Native English-Speaking (NES) and Non-native English-Speaking (NNES) academic writers.

2. THEORETICAL BACKGROUND

Cross-cultural Variation in Written Communication

Regarding the influence of culture on writing, two opposing positions can be witnessed, one focusing on the universality of academic writing (Widdowson, 1979) while the other stressing the cultural differences of textual patterns (Clyne, 1987; Kachru, 1983). Favoring the second view,

Mauranen (2001) asserts that texts are ―cultural products‖ (p, 53) which represent the social interactions occurring in a specific culture. As Kaplan (1966) asserts, specific language cultures exert influence on L2 writing since non-native writers employ L1 writing strategies while writing in L2 (Connor, 1996, 2004; Cumming, 1989) and accordingly, rhetorical structures of writers 'culture are manifested in their L2 texts (Connor, 1996). In the same line, Sharifian (2003, p. 204) proposes the notion cultural conceptualizations defined as ―representations that are distributed across the minds in a cultural group‖ which are constructed by participation in the same cultural context are normally represented in language. Accordingly, there is a two-way interaction between language and cultural conceptualizations since we communicate our conceptualizations through language and on the other side, the linguistic elements we use affect the ways we shape our conceptualizations (Sharifian, 2009). Given that, in utilizing English as an international language, the cultural conceptualizations of individuals from different L1 backgrounds affect the way they write in English as their L2 (Sharifian, 2009). Therefore, writers from different L1 and different cultural conceptualizations may draw on different conventions of L2 writing, which requires in-depth analysis of how culture impacts upon ones 'L2 writing structures (Atkinson, 2003).

In his intercultural rhetoric approach, Connor (2004, p. 293) regards writing as a ―socially situated‖ practice in which the norms of cultural and disciplinary community affect the degree of interaction and collaboration, contextual expectation of audience, and degree of explicitness in the text. Considering the academic writing, the variation in the discourse is attributed to cultural values of writers (Duszak, 1997), which affect different aspects of L2 writing including responsibility over the claims, text organization, evidence presentation, and voice (Steinman, 2003). Moreover, Kaplan (1990) remarks that presenting and supporting evidence to persuade readers are heavily dependable on the culture of the writer. To put it succinctly, ―an academic text reflects the social self-image of the writer and his/her perception of the readership‖ (Duszak, 1997, p. 13). The self-image is witnessed in the writers 'use of stance markers and the readership is reflected in the employment of engagement markers in written texts, which are the foci of the present study.

Hyland's (2005a) Model of Interaction

The present study is premised on the interaction model proposed by Hyland (2005a). Given the fact that the focus of the present study was on stance and engagement markers, and that Hyland 's (2005a) model offers an interactional model which includes both stance and engagement markers and their sub-categories, his model was adopted for the purpose of this study (Figure 1).

According to Hyland (2005a), stance can be defined as ―the ways writers present themselves and convey their judgements, opinions, and commitments‖ (p. 176). Stance markers are the―writer-oriented features of interaction‖ (Hyland, 2008a, p. 9). As outlined by Hyland (2005a), there are four stance markers including hedges, boosters, attitude markers, and self- mentions. Put succinctly by Hyland, hedges (e.g., possible, may) help writers avoid shouldering the responsibility of their claims, boosters (e.g., obvious, surely) represent writers 'degree of certainty about the statements, attitude markers (e.g., interesting, surprisingly) allow writers to express their attitudes toward the information, and self-mentions (e.g., I, we) represent first person pronouns and possessive adjectives in a text which signal the presence of the writer.
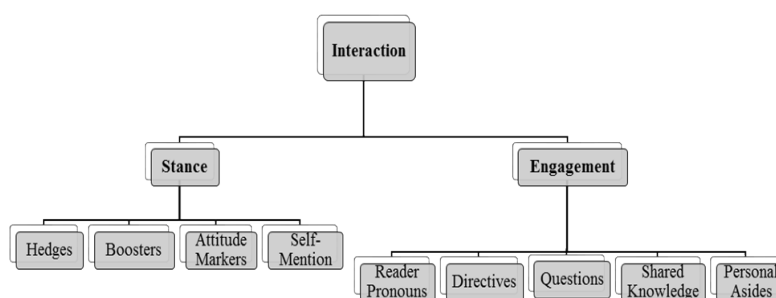
Figure 1: Key resources of academic interaction. (Adopted from Hyland, 2005a, p. 177)

On the other hand, engagement is ―an alignment dimension where writers acknowledge and connect to others, recognizing the presence of their readers‖ (Hyland, 2005a, p. 176). Hyland contended that writers attempt to engage with their readers in five ways including reader pronouns, personal asides, appeals to shared knowledge, directives, and questions. According to him, reader pronouns (e.g., you, reader) are ―the most explicit way that readers are brought into a discourse" (p. 182), personal asides (e.g., by the way) help writers offer explicit comments on the shared information, appeals to shared knowledge (e.g., normally, of course) reflect readers' familiarity and understanding of the common conceptions based on disciplinary conventions, directives (e.g., it's important to, let's) are a form of imposition on readers to do something and to believe something, and questions are renowned for arousing the readers' veritable desire to accompany the writer throughout the text. Below a short review of literature on cross-disciplinary and cross-cultural differences in the use of stance and engagement markers is presented.

3. LITERATURE REVIEW

A steady stream of research on the use of stance and engagement markers in academic writing is witnessed in recent years (Hyland, 2005a; Hyland & Jiang, 2016; Xu & Nesi, 2019) and there have been a number of studies reporting cross-disciplinary and cross-cultural differences in terms of the use of stance and engagement markers in RAs. Hyland (2002a), as one of the prominent figures of the field, focused on the role of first-person pronouns in academic writing. It was found that the use of these pronouns was more evident in RAs of soft disciplines than the ones in hard disciplines. Hyland (2002b), in another study, interviewed expert writers of eight disciplines and analyzed 240 RAs. His results revealed that the writers of hard sciences were more reluctant to make their texts personal than writers of humanities and social sciences. Later, Hyland (2008a) investigated the use of stance and engagement markers in eight disciplines and the results indicated that the humanities and social sciences used these interactional markers more than the science and engineering fields―to adopt more direct positions in their claims and to establish effective interactions with readers.

As for the influence of culture on the use of stance and engagement markers, Martínez (2005) investigated the differences between articles produced by NES writers and RAs produced by NNES Spanish writers in different sections of biology articles. The results suggested that the overall frequency of first- person pronouns used by NES writers was more than the frequencies used by NNES writers. Moreover, Abdollahzadeh (2011) compared the interpersonal metadiscourse in articles written in English by Anglo-American and Iranian academic writers in

the field of applied linguistics. He reported that English writers used more interpersonal markers than did Iranian writers.

In a similar vein, Lafuente-Millán (2014) also examined the frequency of engagement markers in a corpus of business management articles written in two different languages of English and Spanish. It was revealed that Spanish scholars drew more on engagement markers. Çandarlı et al. (2015) also investigated the use of stance markers in English essays written by Turkish and American students. Their results indicated that the frequency of these markers in English essays written by Turkish students was comparably close to the frequency of these markers in essays produced by American students. In a more recent study, Işık- Taş (2018) focused on the frequency of first- person pronouns in the sociology RAs written by Turkish and NES authors published in international and local journals. The results revealed that English and Turkish writers publishing in international journals used first- person pronouns more than Turkish authors publishing in local journals.

In another recent study, Dontcheva- Navratilova (2021) explored the use of engagement markers in linguistic and economic RAs written by Anglophone and Czech authors. The results revealed cross-disciplinary variations with the linguistic RAs using more engagement markers then economics RAs. However, the cultural variation was not highly significant with the two cultures differing significantly only in the use of appeals to shared knowledge which was used more by Czech scholars in order to emphasize the common knowledge and shared conventions between writers and readers. Moreover, there were significant differences in terms of the sub-categories including reader pronouns and types of directives in the locally- published RAs written by Czech scholars and internationally-published RAs authored by Anglophone scholars.

In brief, the majority of studies conducted on cross-cultural and cross-disciplinary variation demonstrate that academic writing is not created in vacuum, but various subject-specific literacies are involved (Hyland, 2002b) and that ―writing is a social act that can occur within particular situations‖ (Hyland, 2009, p. 26). Therefore, the analysis of how and to what degree the use of interactive features in the RAs of authors from different cultural and disciplinary backgrounds differ provides useful insights for novice and experienced scholars and also native and non- native academicians on how interactions occur in academic discourse. However, there have been few studies focusing on all elements of both stance and engagement markers across soft and hard disciplines RAs written by native and Iranian non-native academic writers. Thus, the current study will contribute to this line of enquiry by exploring stance and engagement markers used in the RAs of soft and hard sciences written by native and non-native academic writers. We addressed two research questions in this study:

What are the differences between Iranian academic writers of soft and hard disciplines in their use of stance and engagement markers in English RAs?

How do native speakers of English and Iranian non-native academic writers of English differ in terms of stance and engagement markers in their RAs?

4. THE STUDY

The Corpora

RAs of soft and hard disciplines were selected to compare the patterns of self-representation and interaction across different disciplines in two different cultures. Applied Linguistics, Sociology, and Economics were selected as representatives of soft sciences and Chemical Engineering, Mechanical Engineering, and Biology were selected as representatives of hard sciences (Becher & Trowler, 2001; Hyland, 2002b; Hyland, 2005a). In order to use RAs that are representative of each field, we selected journals suggested as top-tier high-ranking in each

discipline. To ensure the quality of each journal, Scientific Journal Ranking (SJR) was first searched and then the selected journals were checked with specialists from each discipline. The journals selected from different disciplines are listed in Appendix A.

The main corpus of the present study included 120 RAs consisting of 60 papers selected from two different sub-corpora. The first sub-corpus of this study was comprised of RAs written by NES writers. This sub-corpus was designated as _NES sub-corpus '. To find the articles written by NES academics, top-tier journals of each field were browsed and then 10 articles written from 2010 to 2019 were selected from each discipline randomly. In order to determine if the writers were native speakers of English, we drew on some criteria including the authors 'names, institutional affiliations, e-mail addresses provided in RAs, and where they had studied. The articles selected were all written by researchers who had studied and worked in English-speaking countries based on online information. Although it was difficult to identify the nativeness of researchers based on their names and affiliations, it was assumed that the articles were sample models of standard English since they were all published in top-tier journals (Lafuente-Millán, 2014). These criteria were not solely decisive in determining the writers 'originality; however, they helped us make a more accurate decision.

The second sub-corpus of the study belonged to those Iranian writers who had published in international journals. This sub-corpus was designated as _internationally-oriented NNES sub-corpus 'to emphasize the fact that this sub- corpus included nonnative-authored RAs that are published in international English-medium journals. To find RAs written by Iranian writers, we first decided on the journal and then the word Iran or Iranian was searched in the content list of each journal. Then, we checked their names, institutional affiliations, email addresses, and where they had studied. Subsequently, 10 RAs that were published from 2010 to 2019 were selected randomly form each discipline. Selecting this corpus was less demanding since the authors were familiar with Iranian names and their affiliations.

It is worth mentioning that all sections of the articles, except for tables, references, and footnotes, were included in the corpus since authors may draw on interactional markers in different sections of articles such as introduction, methodology, or discussion. The length of each article ranged from 8,000 to 12,000 words depending on the discipline and journal from which it was selected. As shown in Table 1, the final corpus of this study consisted of approximately 1,273,685 words in total, including 747,056 words from RAs in the NES sub-corpus and 526,629 words from RAs in the internationally oriented NNES sub-corpus.

Table 1: Corpus Characteristics

The Sub-corpora	Number of Documents	Number of Words
The NES Sub-corpus	60	747,056
Internationally Oriented NNSE Sub-corpus	60	526,629
The Main Corpus	120	1,273,685

Corpus Analysis Procedure

The two sub-corpora were explored for stance and engagement markers using MAXQDA software which is designed for use in qualitative, quantitative, and mixed-method researches. We first decided to use Nvivo 10, but its preliminary corpus analysis did not yield comprehensive results which would best cover all the purposes of our current research. For instance, the software was not case-sensitive, which was a requirement for the analysis of self-mention (I) or the imperatives as a sub-category of directives. Accordingly, to count the frequency of interactional markers, MAXQDA, which allows for analyzing large amounts of text, was utilized. MAXQDA has several features including the function to set up our own dictionary to investigate the corpus with reference to specific words, using keyword- context function to specify the textual function of selected words, being case-sensitive, being able to recognize phrases—which was helpful in our analysis of word combinations such as *of course, we know that*—and displaying the frequency of selected words visually. Moreover, MAXQDA has an easy-to-use exporting option of the entire data into other statistical software for further statistical analysis.

First, a dictionary including stance and engagement markers was created. The stance and engagement markers found in the literature were extracted. For this purpose, Hyland 's studies were checked completely to find the engagement markers (Hyland & Jiang, 2016) and stance markers (Hyland, 2000, 2005). In general, 291 features including stance and engagement markers (Appendix B) were examined in the two corpora. After creating the dictionary containing stance and engagement markers, RAs were incorporated into MAXQDA to count the frequencies of each marker. Once the corpora were examined, the frequencies of interactional markers in each corpus were calculated. It should be noted that some of the corpus were checked in MAXQDA manually. For instance, the use of imperatives (e.g. look at Table) was checked by one of the researchers to reach a more reliable result. In an instance, the use of phrases such as *'It is known that'* were examined manually to make sure that phrases were accurately taken into account by the software.

In brief, the analysis was done in two separate phases. In the first phase, the frequencies of the markers were calculated across hard and soft disciplines and then in order to investigate whether the differences observed between the two corpora in terms of frequency of stance and engagement markers were significant or not, Chi-square tests for independence were run. In the second phase, the frequencies of the markers in hard and soft science NES sub-corpus were compared with the frequencies found in those of NNES sub-corpus by first calculating the frequencies and then running Chi-square tests to determine the significance of the differences found in the frequencies.

5. RESULTS

The results below report the disciplinary variation among soft and hard disciplines Iranian-authored RAs and cross-cultural variation between English native and Iranian non-native English academic writers in terms of their use of stance and engagement markers. In all the following tables, the frequencies have been normalized to 10,000 words to make cross- corpora comparison possible. As Biber, Conrad, and Reppen (1998) stated ―normalization is a way to adjust raw frequency counts from texts of different lengths so that they can be compared accurately‖ (p. 263). For instance, the total frequency of stance markers in the corpus of NNES Applied Linguistic (103871 words) was 444. However, after normalization per 10000 words, the frequency was 42. Normalization was done as the size of corpora may slightly differ based on the following formula:

Frequency of each marker/ Number of words in the corpus*10000

Disciplinary Variation

In response to the first research question, Chi- square test analyses were run in order to compare Iranian soft and hard disciplines academic writers in their use of stance and engagement markers

in their RAs. Table 2 displays the frequencies of the total number of stance and engagement markers used in these soft and hard disciplines RAs. The results showed that both stance markers (soft=586, hard=286) and engagement markers (soft=136, hard=61) were used in Iranian authored soft disciplines RAs more than hard disciplines RAs.

In order to see whether the differences observed between soft and hard disciplines RAs were significant or not, Chi-square tests were run. The results of Chi-square indicated that stance markers ($\chi^2(1)$ =103.21, p=.000, Cohen 's w=.363 representing a moderate effect size) and engagement markers ($\chi^2(1)$ =28.55, p=.000, Cohen 's w=.380 representing a moderate effect size) were significantly used more in soft disciplines RAs than hard disciplines RAs.

Following that, in order to compare the sub- categories of stance and enagagement markers, Chi-square tests were run separately on each sub-category. Regarding the stance markers, four separate analyses of Chi-square were run in order to compare soft and hard disciplines in their use of self-mentions, attitude markers, hedges, and boosters. Based on the results displayed in Table 3, it can be concluded that all stance markers including self-mentions (soft=108, hard=50), attitude makers (soft=75, hard=38), hedges (soft=256, hard=119) and boosters (soft=147, hard=79) were used more in soft disciplines RAs than hard disciplines Ras written by Iranian academic writers.

Table 2: Frequencies of Stance and Engagement Markers in Iranian Soft vs. Hard Disciplines RAs

		Observed	expected	Residual
Stance Markers	Soft Disciplines	586	436.0	150.0
	Hard Disciplines	286	436.0	-150.0
	Total	872		
Engagement Markers	Soft Disciplines	136	98.5	37.5
	Hard Disciplines	61	98.5	-37.5
	Total	197		

Table 3: Frequencies of Stance Markers in Iranian Soft and Hard Disciplines RAs

Stance Markers		Observed N	Expected N	Residual
Self-mentions	Soft Disciplines	108	79.0	29.0
	Hard Disciplines	50	79.0	-29.0
	Total	158		
Attitude Markers	Soft Disciplines	75	56.5	18.5
	Hard Disciplines	38	56.5	-18.5
	Total	113		
Hedges	Soft Disciplines	256	187.5	68.5
	Hard Disciplines	119	187.5	-68.5
	Total	375		
Boosters	Soft Disciplines	147	113.0	34.0
	Hard Disciplines	79	113.0	-34.0
	Total	226		

Table 4: Frequencies of Engagement Markers in Iranian Soft and Hard Disciplines RAs

Engagement Markers		Observed N	Expected N	Residual
Reader Pronouns	Soft Disciplines	44	29.5	14.5
	Hard Disciplines	15	29.5	-14.5
	Total	59		
Personal Asides	Soft Disciplines	4	3.5	.5
	Hard Disciplines	3	3.5	-.5
	Total	7		
Appeals to Shared knowledge	Soft Disciplines	23	17.0	6.0
	Hard Disciplines	11	17.0	-6.0
	Total	34		
Questions	Soft Disciplines	5	5.0	.0
	Hard Disciplines	0		
	Total	5a		
Directives	Soft Disciplines	60	46.0	14.0
	Hard Disciplines	32	46.0	-14.0
	Total	92		

a. This variable is constant. Chi-square test cannot be performed.

Furthermore, based on Chi-square test analyses, it can be concluded that all stance markers including self-mentions ($\chi^2(1)$=21.29, p=.000, Cohen's w=.367 representing a moderate effect size), attitude makers ($\chi^2(1)$=12.11, p=.001, Cohen's w=.327 representing a moderate effect size), hedges ($\chi^2(1)$=50.05, p=.000, Cohen's w=.365 representing a moderate effect size), and finally boosters ($\chi^2(1)$=20.46, p=.000, Cohen's w=.300 representing a moderate effect size) were significantly used more in soft sciences RAs than hard sciences RAs witten by Iranian academic writers.

Regarding engagement markers, five separate Chi-square test analyses were run in order to compare RAs of soft and hard disciplines in their use of reader pronouns, personal asides, appeals to shared knowledge, questions, and directives. Based on the results displayed in Table 4, it can be concluded that all engagement markers including reader pronouns (soft=44, hard=15), personal asides (soft=4, hard=3), appeals to shared knowledge (soft=23, hard = 11), questions (soft=5, hard=0) and directives (soft=60, hard=32) were more used in soft disciplines RAs than soft discipline RAs written by Iranian academic writers.

Based on the results of Chi-square test analyses for the use of engagement markers in soft and hard disciplines RAs, it can be concluded that reader pronouns were significantly used more in Iranian soft disciplines RAs than hard disciplines RAs ($\chi^2(1)$=14.25, p=.000, Cohen's w=.491 representing a moderate to large effect size), although personal asides were used more in soft disciplines RAs than hard disciplines RAs, the difference was not a significant one ($\chi^2(1)$=.143, p=.705, Cohen's w=.142 representing a weak effect size), appeals to shared knowledge were significantly used more in Iranian soft disciplines RAs than hard disciplines RAs ($\chi^2(1)$=4.23, p=.040, Cohen's w=.352 representing a moderate effect size), the frequency of questions for the hard disciplines RAs was zero hence, Chi-square test was not run for questions, and finally directives were significantly used more in soft disciplines RAs than hard disciplines RAs ($\chi^2(1)$=8.52, p=.004, Cohen's w=.304 representing a moderate effect size).

Cross-cultural Variation

In order to answer the second research question, Chi-square test analyses were run to compare native English and Iranian academic writers in their use of stance and engagement markers in their RAs. Table 5 displays the frequencies of stance and engagement markers used in NES and Iranian NNES corpora. The results showed that stance markers (NES corpus=1577, NNES corpus=872) and engagement markers (NES corpus=385, NNES corpus=197) were used more by NES academic writers than NNES academic writers.

The results of Chi-square indicated that stance markers (χ^2 (1) =202.95, p=.000, Cohen 's w=.287 representing an almost moderate effect size) and engagement markers (χ^2 (1) =60.72, p=.000, Cohen 's w=.323 representing a moderate effect size) were significantly used more in RAs written by NES academic writers than the ones written by NNES writers.

Table 5: Frequencies of Stance and Engagement Markers in NES and NNES RAs

		Observed	expected	Residual
Stance Markers	NES Academic Writers	1577	1224.5	352.5
	NNES Academic Writers	872	1224.5	-352.5
	Total	2449		
Engagement Markers	NES Academic Writers	385	291.0	94.0
	NNES Academic Writers	197	291.0	-94.0
	Total	582		

Table 6: Frequencies of Stance Markers in NES and NNES RAs

Stance Markers		Observed N	Expected N	Residual
Self-Mentions	NES Academic Writers	337	247.5	89.5
	NNES Academic Writers	158	247.5	-89.5
	Total	495		
Attitude Markers	NES Academic Writers	213	163.0	50.0
	NNES Academic Writers	113	163.0	-50.0
	Total	326		
Hedges	NES Academic Writers	648	511.5	136.5
	NNES Academic Writers	375	511.5	-136.5
	Total	1023		
Boosters	NES Academic Writers	379	302.5	76.5
	NNES Academic Writers	226	302.5	-76.5
	Total	605		

Following that, four separate Chi-square test analyses were run in order to compare native and non-native academic writers in their use of self-mentions, attitude markers, hedges and boosters. Based on the results displayed in Table 6, it can be concluded that all stance markers including self-mentions (NES corpus=337, NNES corpus=158) attitude makers (NES corpus=213, NNES corpus=113), hedges (NES corpus=648, NNES corpus=375), and boosters (NES corpus=379, NNES corpus=226) were used more by NES academic writers than NNES academic writers.

Based on the results of Chi-square test analyses for the use of stance markers in native and non-native academic writers' RAs, it can be concluded that all stance markers including self- mentions (χ^2 (1)=64.72, p=.000, Cohen's w=.362 representing a moderate effect size), attitude makers (χ^2 (1)=30.67, p=.000, Cohen's w=.306 representing a moderate effect size), hedges (χ^2 (1)=72.85, p=.000, Cohen's w=.266 representing a weak to moderate effect size), and finally boosters (χ^2 (1)=38.69, p=.000, Cohen's w=.253 representing a weak to moderate effect size) were significantly used more by NES academic writers than NNES academic writers.

Five separate Chi-square tests were run in order to compare native and non-native writers in their use of reader pronouns, personal asides, appeals to shared knowledge, questions and directives. Based on the results displayed in Table 7, it can be concluded that all engagement markers including reader pronouns (NES corpus=126, NNES corpus=59), personal asides (NES corpus=14, NNES corpus=7), appeals to shared knowledge (NES corpus=69, NNES corpus=34), questions (NES corpus=10, NNES corpus=5), and directives (NES corpus=166, NNES corpus=92) were more used more by NES academic writers than NNES academic writers.

The results of Chi-square test analyses for the use of engagement markers in native and non-native academic writers' RAs indicate that reader pronouns were significantly used more by NES academic writers than NNES academic writers (χ^2 (1) =24.26, p=.000, Cohen 's w=.362 representing a moderate to large effect size), although personal asides were used more in NES corpus than NNES corpus, the difference was not a significant one (χ^2 (1)=2.33, p=.127, Cohen's w=.333 representing a moderate effect size), appeals to shared knowledge were significantly used

Table 7: Frequencies of Engagement Markers in NES and NNES RAs

Engagement Markers	Observed N		Expected N	Residual
	NES Academic Writers	126	92.5	33.5
	NNES Academic Writers	59	92.5	-33.5
	Total	185		
	NES Academic Writers	14	10.5	3.5
	NNES Academic Writers	7	10.5	-3.5
	Total	21		
	NES Academic Writers	69	51.5	17.5
Reader Pronouns	NNES Academic Writers	34	51.5	-17.5
	Total	103		
	NES Academic Writers	10	7.5	2.5
	NNES Academic Writers	5	7.5	-2.5
	Total	15		
	NES Academic Writers	166	129.0	37.0
	NNES Academic Writers	92	129.0	-37.0
	Total	258		

more by NES academic writers than NNES academic writers (χ^2 (1)=11.89, p=.001, Cohen's w=.339 representing a moderate effect size), although questions were used more in NES corpus than NNES corpus , the difference was not a significant one (χ^2 (1)=1.66, p=.197, Cohen's w=.332 representing a moderate effect size), and directives were significantly used more by NES academic writers than NNES academic writers (χ^2 (1)=21.55, p=.000, Cohen's w=.290 representing an almost moderate effect size).

6. DISCUSSION

Disciplinary Variation

The findings in the current corpus analysis research demonstrated that soft disciplines RAs had more stance and engagement markers than the ones in hard disciplines. In this sense, linguistic elements found in RAs are constrained by the conventions of the discourse community for whom they have been written. The findings of the present study are in line with Hyland (2005a) who believed that ―rhetorical practices are inextricably related to the purposes of the disciplines‖ (p. 187) which implies that patterns of representing one's self and engaging the readers in RAs is a discipline-related issue with most of the time writers of soft sciences and humanities representing themselves more directly and explicitly than the writers of science and engineering fields. As Hyland (2018) states, disciplines ―have a very real existence for those who work and study in them‖ (p. 29). Accordingly, in hard sciences knowledge construction is based on the empirical evidence and it is also ―cumulative and tightly structured‖ (Hyland, 2005a, p. 188). As Hyland (2018) puts it, in hard sciences the discourse community puts emphasis on the research itself and neglects the importance of researcher; hence, the role of author is downplayed in RAs of hard disciplines to emphasize the importance of the knowledge and content which are conveyed throughout the text. However, in humanities and soft sciences, writers have a wide variety of readers so ―personal credibility, and explicitly getting behind arguments‖ are more important in having a persuasive discourse (Hyland, 2005a, p. 188). In this sense, writers are expected to express ideas in a way which allows different interpretations from different perspectives due to the open nature of soft disciplines (Becher & Trowler, 2001). Moreover, due to the flexible nature of soft sciences, academic writers tend to use more stance markers (Abdi, 2011), while hard science academic writers present the arguments more rigidly due to the factual nature of their disciplines (Becher & Trowler, 2001).

Regarding the frequencies of each stance marker, hedges were the most frequent marker in both hard and soft sciences; however, RAs of soft disciplines enjoyed higher frequencies of hedges in comparison to hard disciplines RAs. This is comparable with Hyland (1998, 2005a, 2008a) who indicated that hedges were used more frequently in soft sciences to show lower degrees of certainty toward the claims and this implies the fact that academicians attempt to present facts and their claims with caution. Following hedges, RAs of both disciplines enjoyed considerable rate of boosters. Boosters were found more in soft disciplines RAs than their counterpart hard discipline RAs and this is in line with Vázquez Orta and Giner (2009) in that they observed more boosters in marketing RAs in comparison to the ones in mechanical engineering and biology. In addition, this is consistent with Hyland (2008a) who reported more usage of boosting elements in RAs of soft disciplines. Indeed, sociological aspects of each discipline affect the use of these interactional markers in a way that RAs of soft sciences present more imprecise results requiring stronger and clearer supports from the writers which results in the use of boosters. However, authors of hard disciplines report the results more precisely; hence, they do not need to draw on boosters to

convince their readers. However, there are a considerable number of boosters used in hard sciences to ensure the readers of the credibility of the results.

Results also revealed that self-mentions in soft disciplines RAs are used significantly more than the ones in hard disciplines. This may underline the fact that academicians of soft sciences prefer to infuse their voice into the text while the academicians of hard sciences avoid doing so and try to efface their voice from RAs. The findings of the present study are supported by Hyland (2001) in that he argues that academicians of hard sciences refrain from presenting themselves and instead put emphasis on the objectivity of the results by believing that the data should be reported irrespective of the subjective view points of the author. However, higher rates of self-mentions in soft sciences offer a stronger expression of authors 'presence to present the less solid results with more responsibility and confirmability. This is also consistent with Hyland (2018) who believes academicians of soft disciplines are encouraged —to present their own voice 'and display a personal perspective, suitably supported with data and intertextual evidence" (p. 174). However, writers of hard sciences avoid using self-mentions to —highlight the phenomena under study, the replicability of research activities, and the generality of the findings‖ (Hyland, 2008a, p. 17).

Attitude markers were the least frequently- used markers in both disciplines; however, their frequency was higher in soft disciplines RAs and this is in line with Hyland (1998) who believes that lower frequencies of attitude markers in hard sciences reflect writers 'unwillingness to wield their authority in stating the claims. In the same vein, attitude markers in hard science RAs are used to answer the questions that readers may have in mind, based on background knowledge, to satisfy readers 'needs; however, soft disciplines draw on a wide range of attitude markers to reflect their affection and attitudes toward their claims irrespective of readers 'background knowledge.

In addition to constructing authority and credibility by using stance markers, —writers are able to either highlight or downplay the presence of their readers in the text‖ by using engagement markers (Hyland, 2008a, p. 17). It was revealed that RAs of soft sciences employ more engagement markers than hard sciences RAs which is in line with Hyland (2005a, 2008a). Moreover, it was revealed that directives were the most frequently-used markers in both hard and soft sciences RAs. It is shown that the only frequent interactive feature used in engineering and sciences were directives and this may imply that hard sciences RAs are mainly short and concise; thus, using imperatives provides writers with the economy of words (Hyland, 2005a). Furthermore, it can be inferred that employing directives in RAs is a useful strategy that writers draw on to build rapport with readers either in soft or hard disciplines. This is in accordance with Hyland's (2005c) study which revealed that the total number of directives was higher than other engagement markers in all eight disciplines of his study.

Following directives, reader pronouns had the highest frequency. According to Hyland (2005a), authors use reader pronouns —to appeal to scholarly solidarity, presupposing a set of mutual, discipline-identifying understandings linking writer and reader‖ (p. 188). This underlines the fact that reader pronouns are the best devices through which writers can connect to their readers and allow readers to accompany them in each section of the RAs. However, writers of hard sciences avoid using reader pronouns in their texts and it is supported by Hyland and Jiang (2016) who observed that reader pronouns are almost never used in engineering and sciences. Furthermore, Hyland (2005c) contends that knowledge construction in soft sciences is tremendously interpretive thus —proofs must appeal to the reader 's willingness to follow and accept the persuasiveness of the discourse‖ (p. 370) and given that, authors of soft sciences draw on reader pronouns to persuade readers to accompany them throughout the texts.

Furthermore, questions were mainly found in soft disciplines RAs and this finding is supported by Hyland (2005a). Questions used in soft sciences RAs reflect the existence of audience to draw

their attention to writers 'statements and to persuade readers to answer the statements in advance and make themselves compatible with the arguments which is going to be followed. Hyland (2008a) also noted that questions are ―almost exclusively confined to the soft fields‖ (p. 18). Personal asides were the least-frequent markers in both disciplines as is the case for Hyland (2005c). Differences regarding personal asides in this study were not significant. However, personal asides and appeals to shared knowledge were used more frequently in humanities than hard sciences and this may imply the fact that in soft sciences, writers depend on shared knowledge and common beliefs to express their ideas to show that writers and readers share common conceptions. As Hyland (2005c) believes, appeals to shared knowledge and personal asides are useful devices in ―emphasizing shared goals and drawing the reader into the discourse as a fellow disciplinary member‖ (p. 368).

Cross-cultural Variation

With regards to cultural variation, the differences between NES and Iranian NNES academic writers were reported and it was revealed that using interactional markers in writing is affected by one 's culture. It is believed that the conventions of academic writing differ from culture to culture (Clyne, 1991) since writers have different cultural conceptualizations which affect their use of linguistic resources (Sharifian, 2009). The two groups of writers were statistically different from each other in terms of both stance and engagement markers in a way that NES writers used more interactional markers than their counterpart NNES Iranian academics. This implies that culture can profoundly influence the linguistic choices of writers and as Yakhontova (2006) stated, ―culture specific differences‖ are evident in writing styles of different academicians. The findings are supported by Martínez (2005) who showed that the overall frequency of first-person pronouns in NES corpus was higher than what was found in NNES Spanish corpus. This shows that native academic writers try to exert power in their texts to prove the originality of their works to the discourse community so that they can attain acceptance and recognition.

The importance of originality and focus on the individual requires having a voice in academic writing (Atkinson, 2003) and western cultures consider patterns of self-representation as a sign of ―strength, confidence, and individuality‖ in their written texts (Steinman 2003, p.83). This finding also resonates well with Dontcheva-Navratilova (2021) whose results revealed that in comparison to RAs authored by Anglophone writers, RAs by Czech scholars used fewer first-person plural pronoun in the field of Economics, which indicated the writer-oriented approach of Czech scholars and less degrees of writer-reader interactions. The findings imply that English authors not only draw on more interactional markers, but also construct more reader-oriented texts to help readers perceive the texts better. Emphasis on effective communication with readers in written discourse is taught to English students in their school systems (Dahl, 2004); accordingly, English writers construct more reader-friendly approaches to writing and establish higher rates of interaction with their audience (Thompson, 2001).

Moreover, the findings are supported by Abdollahzadeh (2011) who showed that English writers used more interpersonal markers than did Iranian writers. This implies that English writers aim to establish rapport with their audiences so that they can invariably convince their readers of the credibility of their claims. A reader-oriented approach may be adopted by writers drawing on monologic and dialogic views (Cmejrková & Danes, 1997). Western cultures favor the dialogic stance which focuses on establishing rapport with readers who can be actively engaged in the text, while Iranian scholars adopt the monologic stance stressing the rigid production of texts with an emphasis on the truth of the data. This shred of evidence is also in line with Dontcheva-Navratilova (2021) whose results indicated more frequent use of directives in the results and discussion

sections of RAs authored by Anglophone academicians than their Czech counterparts, which may imply Anglophone scholars 'intentions to interact and engage with a more diverse audience.

English native speaker academicians are familiar with a wide range of lexicogrammatical features of English (Xu & Nesi, 2019); therefore, by drawing on various resources they can build credible identity and construct rapport with their audience in order to ―gain acceptance for their claims through a balanced demonstration of deference, humility, respect, attitudinal and assertive language to persuade readers about the validity of their arguments‖ (Abdollahzadeh, 2011, p. 292). Furthermore, this can also be associated with the fact that western cultures are more inclined toward individualism than Eastern cultures (Crismore, Markkanen, & Steffensen, 1993; Hofstede, 1997); hence, they employ more interactional markers to ascertain their claims more explicitly in order to represent an authoritative stance toward their claims to bolster the effect of their arguments. However, many writers from Asian countries tend to exert less authorship in their texts and also write less explicitly (Hinkel, 2002). As indicated, western cultures are individualistic in a sense that they favor direct and explicit presentation of materials while oriental societies stress the collectiveness and connectedness which accordingly make the text less assertive and the claims more flexible (Duszak, 1997).

In addition, the results imply that Iranian authors are adhering to the traditional views of objective academic writing in which authors refrain from infusing their own voice and building rapport with readers since accepting the traditional beliefs without trying to change or challenge them is a part of Iranian culture (Abdollahzadeh, 2011). This is also corroborated by Markus and Kitayama 's (1991) view of self in different cultures who assert that there are two views of independence and interdependence with the former stressing ―the inherent separateness of distinct persons‖ (p. 226) and the latter focusing on connectedness of individuals to one another. They assert that western cultures favor the independent view while individuals from interdependent cultures try to represent the traditions of their culture rather than writing to present themselves as individual selves. Thus, in the case of Iranian scholars, they are adopting the interdependent view in which traditional beliefs of writing are deemed more appropriate.

7. CONCLUSIONS

In the present study, a corpus-based approach was adopted to investigate how authors of various disciplines and from different cultures employ stance and engagement markers in their RAs. The findings of the present study reflected the differences in the distribution of these markers in the two corpora. Regarding disciplinary variation, significant differences were observed among soft and hard disciplines RAs in that writers of soft sciences drew on more interactional markers than their counterpart hard sciences 'writers. In exploring the effect of culture, it was revealed that culture plays a significant role and the results implied that NES academic writers tended to use more stance and engagement markers in their RAs than did NNES Iranian academics.

The concepts of stance and engagement markers represent the ways authors use community-tied and culture-bound linguistic devices to express their authority in texts and also to interact with their readers. Perhaps the most notable implication of the present study purports to reflect how the use of the linguistic devices in academia is dependent on disciplinary discourse to which we belong and the culture we come from to represent ourselves as successful linguists, sociologists, engineers, biologists, etc. (Çandarlı et al., 2015; Hyland, 2008a). Put simply, corpus-based approaches focusing on stance and engagement markers in RAs can help academicians understand how these interactional markers are utilized by NES and NNES academic writers across different fields of study.

With regard to the pedagogical implications of the current study, this study should make academic writers conscious of their own presence in their texts and the interaction they can build

with readers based on the use of stance and engagement markers. We believe that the members of a particular discourse community should be aware of so many discipline-specific norms, preferences, and also the ways in which presence and interaction are affected by the disciplinary conventions and cultural contexts. Investigating cultural patterns of academic writing helps academicians get cognizant and conscious of both cross-cultural and cross- disciplinary variations (Steinman, 2003) and will accordingly equip writers with the knowledge to write more professionally. By making students and academics cognizant of the appropriate use of stance and engagement markers, we not only improve their understanding of disciplinary and cultural conventions but also prepare them for projecting their own arguments in their community of practice.

Pedagogically speaking, the teaching of discourse markers should be more explicitly included in university courses. Explicit instructions can be made possible through different EAP courses and also advanced writing classes in which novice scholars get familiar with the use of these markers through different reading and writing tasks in different subject areas. We also suggest that EAP instructions in Iranian context would help researchers understand the differences between their L1 and L2 rhetorical conventions.

As any other research projects, this research has also some methodological limitations as corpus-based studies do not offer examination of the hidden processes that writers employ during the use of stance and engagement markers. However, the mentioned limitation can be solved by applying more qualitative approaches such as narrative enquiries, discourse analysis, ethnographic approaches, and interviews which shed light on the perspectives of native and non- native academicians about the motives behind the inclusion of these linguistic elements. Another suggestion for future research concerns the proficiency level of the sample selected. Future studies should be directed at incorporating RAs from both students and experts in different disciplines to explore how proficiency level would affect the use of discourse markers.

References

Abdi, R. (2011). Metadiscourse strategies in research articles: A study of the differences across subsections. *Journal of Teaching Language Skills, 3*(1), 1-16. https://doi.org/10.22099/jtls.2012.391

Abdollahzadeh, E. (2011). Poring over the findings: Interpersonal authorial engagement in applied linguistics papers. *Journal of Pragmatics, 43*(1), 288-297. https://doi.org/10.1016/j.pragma.2010.07.019

Atkinson, D. (2003). Writing and culture in the post-process era. *Journal of Second Language Writing, 12*(1), 49-63. https://doi.org/10.1016/S1060-3743(02)00126-1

Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the cultures of disciplines* (2nd ed.). Open University Press/SRHE.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Çandarlı, D., Bayyurt, Y., & Marti, L. (2015). Authorial presence in L1 and L2 novice academic writing: Cross-linguistic and cross-cultural perspectives. *Journal of English for Academic Purposes, 20*, 192-202. https://doi.org/10.1016/j.jeap.2015.10.001

Clyne, M. (1987). Cultural differences in the organization of academic texts: English and German. *Journal of Pragmatics, 11*(2), 211-241. https://doi.org/10.1016/0378-2166(87)90196-2(

Clyne, M. (1991). The socio-cultural dimension: The dilemma of the German-speaking scholar. In H. Shroder (Ed.), *Subject-oriented texts, language for special purposes and text theory* (pp. 49–68). De Gruyter.

Cmejrkovä, C., & Danes, F. (1997). Academic writing and cultural identity: The case of Czech academic writing. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 41-62). Mouton de Gruyter.

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second language writing*. Cambridge University Press.

Connor, U. (2004). Intercultural rhetoric research: Beyond texts. *Journal of English for Academic Purposes, 3*(4), 291-304. https://doi.org/10.1016/j.jeap.2004.07.003.

Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication, 10*(1), 39-71. https://doi.org/10.1177/0741088393010001002

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning, 39*, 81-141. https://doi.org/10.1111/j.1467-1770.1989.tb00592.x

Dahl, T. (2004). Textual metadiscourse in research articles: A marker of national culture or of academic discipline? *Journal of Pragmatics, 36*(10), 1807-1825.

Dontcheva-Navratilova, O. (2021). Engaging with the reader in research articles in English: Variation across disciplines and linguacultural backgrounds. *English for Specific Purposes, 63*, 18-32.

Duszak, A. (1997). Cross-cultural academic communication: A discourse-community view. In A. Duszak (Ed.), *Culture and styles of academic discourse* (pp. 11-40). Mouton de Gruyter.

Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*. Lawrence Erlbaum Associates.

Hofstede, G. (1997). *Cultures and organizations: Software of the mind*. McGraw-Hill.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics, 30*, 437-455. https://doi.org/10.1016/S0378-2166(98)00009-5

Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes, 20*(3), 207-226. https://doi.org/10.1016/S0889-4906(00)00012-0

Hyland, K. (2002a). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*(8), 1091-1112. https://doi.org/10.1016/S0378-2166(02)00035-8(02

Hyland, K. (2002b). Options of identity in academic writing. *ELT Journal, 56*(4), 351-358. https://doi.org/10.1093/elt/56.4.351.

Hyland, K. (2005a). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173-192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2005b). *Metadiscourse: Exploring interaction in writing*. Continuum.

Hyland, K. (2005c). Representing readers in writing: Student and expert practices. *Linguistics and Education, 16*(4), 363–377. https://doi.org/10.1016/j.linged.2006.05.002

Hyland, K. (2008a). Disciplinary voices: Interactions in research writing. *English Text Construction, 1*(1), 5-22. https://doi.org/10.1075/etc.1.1.03hyl

Hyland, K. (2008b). Persuasion, interaction and the construction of knowledge: Representing self and others in research writing. *IJES, 8*(2), 1-23. https://doi.org/10.6018/ijes.8.2.49151.

Hyland, K. (2009). *Teaching and researching writing*. Routledge.

Hyland, K. (2018). *The essential Hyland*. Bloomsbury.

Hyland, K., & Jiang, F. (2016). We must conclude that: A diachronic study of academic engagement. *Journal of English for Academic Purposes, 24*, 29-42. https://doi.org/10.1016/j.jeap.2016.09.003

Işık-Taş, E. E. (2018). Authorial identity in Turkish language and English language research articles in Sociology: The role of publication context in academic writers' discourse choices. *English for Specific Purposes, 49*, 26-38. https://doi.org/10.1016/j.esp.2017.10.003

Kachru, Y. (1983). Linguistics and written discourse in particular languages: Contrastive studies: English and Hindi. *Annual Review of Applied Linguistics, 3*, 50-77.

Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural communication. *Language Learning, XVI*, 1-20.

Kaplan, R. B. (1976). A further note on contrastive rhetoric. *Communication Quarterly, 24*(2), 12-19.

Kaplan, R. B. (1990). Writing in a multilingual/multicultural context: What's contrastive about contrastive rhetoric? *Writing Instructor, 10*(1), 7-18.

Lafuente- Millán, E. (2014). Reader engagement across cultures, languages and contexts of publication in business research articles. *International Journal of Applied Linguistics, 24*(2), 201-223. https://doi.org/10.1111/ijal.12019

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224. https://doi.org/10.1037/0033-295X.98.2.224

Martínez, I. A. (2005). Native and non-native writers' use of first-person pronouns in the different sections of biology research articles in English. *Journal of Second Language Writing, 14*(3), 174-190. https://doi.org/10.1016/j.jslw.2005.06.001.

Mauranen, A. (2001). Descriptions or explanations? Some methodological issues in contrastive rhetoric. In M. Hewings (Ed.), *Academic Writing in Context* (pp. 43-54). University of Birmingham Press.

McGrath, L., & Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices. *English for Specific Purposes, 31*(3), 161-173. https://doi.org/10.1016/j.esp.2011.11.002

Sharifian, F. (2003). On cultural conceptualizations. *Journal of Cognition and Culture, 3*(3), 187-207. https://doi.org/10.1163/156853703322336625

Sharifian, F. (2009). Cultural conceptualizations in English as an international language. In F. Sharifian (Ed.), *English as an international language: Perspectives and pedagogical issues* (pp. 242-253). Multilingual Matters.

Shaw, P. (2003). Evaluation and promotion across languages. *Journal of English for Academic Purposes, 2*(4), 343-357. https://doi.org/10.1016/S1475-1585(03)00050-X

Steinman, L. (2003). Cultural collisions in L2 academic writing. *TESL Canada Journal,* 80-91. https://doi.org/10.18806/tesl.v20i2.950

Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics, 22*(1), 58-78. https://doi.org/10.1093/applin/22.1.58.

Vázquez Orta, I., & Giner, D. (2009). Writing with conviction: The use of boosters in modelling persuasion in academic discourses. *Revista Alicantina de Estudios Ingleses, 22*, 219-237. https://doi.org/10.14198/raei.2009.22.14.

Widdowson, H. G. (1979). *Aspects of language teaching*. Oxford University Press.

Xu, X., & Nesi, H. (2019). Differences in engagement: A comparison of the strategies used by British and Chinese research article writers. *Journal of English for Academic Purposes, 38*, 121-134. https://doi.org/10.1016/j.jeap.2019.02.003

Yakhontova, T. (2006). Cultural and disciplinary variation in academic discourse: The issue of influencing factors. *Journal of English for Academic Purposes, 5*(2), 153-167. https://doi.org/10.1016/j.jeap.2006.03.002

Yang, W. (2014). Stance and engagement: A corpus-based analysis of academic spoken discourse across science domains. *LSP Journal - Language for Special Purposes, Professional Communication, Knowledge Management and Cognition, 5*(1), 62-78.

The Effects of Critical Pedagogy and Task-Based Language Teaching on Storytelling and Oral Proficiency: A Comparative Study

Fateme Alikahi[1,*] ID & Gholam Reza Kiany[1]

[1] Tarbiat Modares University, Tehran, Iran

Abstract

Maximizing learning opportunities has long intrigued teaching practitioners and researchers. Therefore, a lot of studies have been conducted on different instructional methodologies to help language learners with utilizing their learning chances. The present comparative study aims at investigating the effects of two instructional procedures, namely Critical Pedagogy (CP) Based Teaching and Task- Based Language Teaching (TBLT) on oral proficiency and storytelling skills of the participants. Thus, 30 Iranian male EFL learners, who were all rated at A1 level on Quick Placement Test (Cambridge, 2001), were selected. All of the learners who aged from 13 to 15 years old were randomly assigned into two research groups of CP and TBLT through convenience and purposive sampling. Pretest/posttest research design was conducted to trace any significant effects of two instructional procedures on storytelling skills and oral proficiency level of the learners before and after 12 sessions of treatment in each research group. Although the results of independent t-tests and effect size indicated impact of each instruction type on learners' storytelling skills and oral proficiency level, the results failed to show any significant difference between Critical Pedagogy Based Teaching and Task- Based Language Teaching. Implications of the study were also discussed further and suggestions for more research were proposed.

1. INTRODUCTION

Research has focused on pedagogical merits of different instructional procedures (R. Ellis, Skehan, Li, Shintani, & Lambert, 2019) and much research has been advocated to comparing more recent methods and traditional ones. The latter has been criticized for not paying enough attention to the active role of students, learner's autonomy, individual development as well as their failure in assisting learners in acquiring critical thinking and transferring what they learn in the context of classroom to the real world (Aghchelu, 2013; Assalahi, 2013; Takala, 2016). Recent learner-centered approaches and their impact on students' scores have intrigued teaching practitioners. Therefore, tasked based instruction and recently critical pedagogy have received profound attention. Task based instruction is an approach where learners comprehend, manipulate, and

* **Corresponding Author:** Fateme Alikahi, English Department, Faculty of Humanities, Tarbiat Modares University, Tehran, Iran.
Email: f.alikahi@yahoo.com

produce language while they focus on meaning yet not ignore form (Mohamadi & Rahimpour, 2018). Critical pedagogy (CP) is about "social transformation" (Akbari, 2008). CP aims at making students conscious of what is happening around them, empowering them to express their voice (Bain, 2010) and encouraging them to be critical to "power relations"(McLaren, 1992). CP helps students to bring social justice to their life and society (Crookes, 2012). According toKumaravadivelu (2006), CP is about relating education to the social, cultural, and political issues, not just limiting it to the phonological, syntactic, and pragmatic domains of language.

Previous research on task-based instruction and critical pedagogy, though important and timely, is limited. The research mainly has covered how task types and their implementation have affected student language quality such as fluency (Albino, 2017), accuracy (Van de Guchte, Rijlaarsdam, Braaksma, & Bimmel, 2019) and complexity (Zenouzagh, 2020). Research has also investigated how TBLT and CP affect the psycholinguistic aspects such as student risk taking, action taking (Vogel, Rudolf, & Scherbaum, 2020), willingness to communicative (Cutrone & Beh, 2018), student attitudes (Zhang & Zhang, 2021) and student engagement (Mohamadi, 2017b). Furthermore, studies have displayed how CP proposes that L2 writers may become more socially and linguistically responsible individuals as they write and reflect on their own experiences with language differences (Britton & Leonard, 2020). They have also focused on teacher and student perception of CP advocating their support (Atai & Moradi, 2016). With respect to critical pedagogy, much of research is concerned with theory and it calls for more practice studies which initiate and sustain discussion dialogues through problem posing to create social transformation (Akbari, 2008; S. Sadeghi, 2008). Considering the incomplete picture of research investigating the effects of CP in practice displayed above and having noticed the cornerstone shared between task-based instruction and CP which is reflection, the researchers intended to investigate the potential of the two in fostering oral proficiency and storytelling in a comparative sense. The researchers considered language proficiency and storytelling skills as the dependent variables since both have been reported to play essential roles in guiding students to scrutinize the information they are exposed to and develop their own understanding and ideas (Shin & Crandall, 2014).

2. LITERATURE REVIEW

Maximizing learning opportunities has long intrigued teaching researchers and practitioners. To this end, research has been conducted on comparing different instructional methodologies among which CP and task- based instruction received profound attention. Although CP has been proposed for three decades, it has not been employed in the educational settings (Mohammadi, Motallebzadeh, & Ashraf, 2014). Naiditch (2016) made a dichotomy of students by making a distinction between subjects and objects directing their learning process. In contrast to more conservative approaches encouraging rote learning, CP sees learning more than mere reception of information and considers learning as a development resulted and originated by students' internal resources. It is associated with a learner-centered curriculum that promotes active, inquiry-based learning and critical co- investigation. Likewise, the task-based instruction focuses on a learner-centered curriculum, which engages learners in real life authentic tasks. Researchers have conducted several studies on each of the abovementioned methodologies and how they influence learning in terms of achievement. However, it seems that applying CP in young learners' classroom with low language proficiency is of a challenge (Abednia, 2015) and more studies aiming to compare these two instructional methodologies with respect to oral proficiency and storytelling are needed.

Task Based Instruction

Task Based Instruction (TBI) is method which considers tasks and completes them as the core of instruction which encourages students' construct of language through interaction and negotiation of meaning. Students also attend to form, active comprehension, manipulation and production of language and promoting students' activeness (R. Ellis et al., 2019). Advocates of Social Constructivism such as Vygotsky prioritizes meaning and emphasizes group work and cooperation believing that through interaction and mediation, children can optimize their learning potential (Brown, 2014). Task-based language teaching is a vibrant area of second language acquisition research (Ashraf Ganjouee, Ghonsooly, & Hosseini Fatemi, 2018). It engages learners with real life authentic tasks which are meaningful to the learners and can activate natural, practical and functional use of language for meaningful purposes (Lin, 2009). TBLT also does not separate four language skills; rather, it integrates them. Tasks implementation in class consists of three main stages. First, pre- task which means preparing the students with required linguistics knowledge and psycholinguistic conditions by the teacher. Second, task- cycle when the students are engaged in actual task accomplishment through negotiation of meaning and interaction and their discourse management, and finally there is a post- task stage when the teacher raises students' consciousness towards language forms (Ansarin & Mohamadi, 2013a, 2013b; Mohamadi, 2015, 2017a) and requires them to produce the language. It is also worth mentioning that a range of different communicative task types, namely information gap tasks, jigsaw, opinion exchange tasks, discussion tasks, role plays, problem solving tasks and their implementation offer a great flexibility and opportunities in directing student learning (R. Ellis et al., 2019).

Critical Pedagogy

There are different terms for Critical Pedagogy such as critical theory, critical literacy, critical applied linguistics and critical language awareness (Hawkins & Norton, 2009). It is said that Critical Pedagogy roots in Marxism (Gruenewald, 2003). To know more about Critical Pedagogy, we start with Paulo Freire's work which first brought the concept of CP to the classroom (Boegeman, 2013)**.** Freire was a teacher in Brazil who worked with slum dwellers. He focused on students' problems in their lives as their literacy program (Larsen-Freeman & Anderson, 2013). Freire has been a pioneer in CP and his role and what he has done in this field are noticeable. Posing real problems, finding solutions for them and dialogical method not only helped students develop literacy but also enabled them to act upon their lives and change their status quo. CP tries to move away from teacher fronted classroom and "instill in students a critical mind-set to become agents of change (Mahmoodarabi & Khodabakhsh, 2015).

Many studies have attempted to suggest practical ways for fostering critical pedagogy. For example, Ghahremani-Ghajar and Mirhosseini (2005) conducted a study on critical pedagogy and journal writing. The results of data analysis on informal written interviews and 6oo journal entries indicated that journal writing can empower students to express their feelings and ideas. Likewise, Ro's (2015) study of implementing CP through drama in an Intensive English Program at intermediate level. Topics like racism, lookism, gender inequality, income inequality and Big Brother were worked on. Results of participants' attitudes, a questionnaire semi-structured interviews, self-reflection papers and students' diaries revealed that CP classroom increased critical perspectives and opportunities for L2 learning through critical dialogue.

Storytelling

According to Wright (1995) "the world is full of storytellers". Storytelling is also important for educational purposes. They provide opportunities for people to share their feelings and build up their confidence and develop their social and emotional understanding (G. Ellis & Brewster, 2014).

Through stories, people become informed of their rights and "the values of democracy and harmony" (G. Ellis & Brewster, 2014). Stories aid critical discussions and help the audience see the world from the characters' points of view and thus comprehend the world diversities better. Smith Byron (2011) probed how storytelling can promote social changes. Participants of this qualitative study which used a grounded theory were ten postsecondary educators. An educational agenda based on standard learning and critical thinking skills challenged the concepts of democratic education. As a result of this study, it was depicted that storytelling was a tool for critique and the center of loving praxis.

Oral proficiency

One of the ultimate goals of EFL is to improve oral communication skills as the output of learning. Student success is evaluated in terms their ability to speak fluently using the target language (Adiantika & Purnomo, 2018; Hasan, 2014). Speaking skills are instruments to facilitate a communication process which covers several aspects of language i.e. Grammar, vocabulary, and pronunciation (Oradee, 2012). Speaking skills (speaking proficiency) are generally used as an indicator of students' competence in learning English (Richards, 2008).

Present study

Considering the incomplete picture displayed above, the researchers intended to investigate the potential of the two instructions in fostering oral proficiency and storytelling in a comparative sense. Oral language proficiency and storytelling skills are the dependent variables since both have been reported to have essential roles in guiding students to scrutinize the information they are exposed to and develop their own understanding (Shin & Crandall, 2014). For this purpose, the following research questions were set to find answers:

1. Is there any significant difference between Critical Pedagogy-Based Teaching and Task- Based Language Teaching (TBLT) in terms of improving learners' oral proficiency?
2. Is there any significant difference between Critical Pedagogy-Based Teaching and Task- Based Language Teaching (TBLT) in terms of improving learners' storytelling skills?

3. METHOD

Participants

From 70 invited participants, 30 Iranian male EFL learners atA1 level rated on Quick Placement Test with the age range of 13-15 from Pakdasht located to the east of Tehran were randomly selected. Their L1 was Persian. The participants were randomly assigned to critical pedagogy group (N=15) and Task-based language group (N= 15). Informed consent was received from all participants.

Instrument

To measure dependent variables, namely oral proficiency and storytelling skills which the former is based on three criteria: Pronunciation, interaction and vocabulary (Cambridge, 2018) and the latter includes setting, initiating events, internal response, attempt and consequence (Merritt & Liles, 1987) were used.

Quick placement test.

To check the level of students' language proficiency and make sure of their homogeneity, a Quick Placement Test (Cambridge, 2001) was given to the participants. The Quick Placement Test (QPT) consists of 60 multiple choice questions and provides teachers with a reliable and efficient

means of placing students at the start of a course. The tests have been calibrated against the levels system provided by the Common European Framework of Reference for Languages, Learning, Teaching, Assessment (commonly known as the CEFR), which has been adopted by the Association of Language Testers in Europe (ALTE). This test, QPT, can clearly and reliably identify any learner's CEFR level (from A1 to C2 CEFR scale) and also can provide a score which shows where the learner is within that band. According to the test, those who answer 17 or fewer questions correctly are at the A1 level. In this study, those who scored between 5 and 17 correct answers were chosen and those who did not belong to this range were excluded.

Cambridge young learners English test.

To measure Oral Proficiency two versions of Cambridge Young Learners English in Pre- and post-test were used. There are three levels to this test: starters (pre- A1), movers (A1), flyers (A2). Both pre-test (Cambridge, 2014) and post-test (Cambridge, 2013) were at Movers level including four parts. In part one, there are two similar pictures with few differences. The candidate is asked to identify the differences between the two photos. In part two, pictures of a story are given to the candidate and he/she should describe each picture in turn and complete the story. Then four sets of photos are given to the candidate and he/she should find the picture, which is different from others. Finally, the interlocutor asks the candidate some open-ended questions and he/she should answer them. Candidates were scored from zero to five for their range of vocabulary, pronunciation and interaction. The intra-rater repeatability resulted from intraclass correlation was high for oral proficiency in pre-test and post-test being 0.994 and 0.988, respectively.

Story grammar component criteria (Merritt & Liles, 1987).

The stories used in pre-test, the tortoise and the hare, and post- test, the jackal and the crow, needed to be analyzed in terms of containing story grammar components. There are various adapted version of Stein and Glenn's story components (Stein & Glenn, 1975). A story grammar is more like a hierarchy including setting and episode categorized initiating events, external response, internal response, attempt, consequence and reaction (Stein & Glenn, 1975). In this study, we used the adapted version which was developed by (Merritt & Liles, 1987) with a little modification, excluding reaction to decide which components each story contains (see Appendices A and B). Two raters analyzed the stories independently based on the Merritt and Liles' criteria for story grammar components. In case of inconsistency, a third person was asked to comment and the answer that two raters out of three agreed on was chosen. Furthermore, having used the same model to calculate the intra-rater repeatability for story grammar components, intraclass coefficient was 0.993 in pre-test and 0.996 in post-test.

Story scoring criteria (Merritt & Liles, 1987).

Students' stories told in pre-test and post-test were transcribed verbatim in addition to the criteria introduced in Appendix A, each statement was included in the analysis if they met the criteria found in Appendix B. There are seven conditions introduced that needed to be met so that statements could be included in the analysis. There are also ten more conditions that clarify which sentences must be discarded. The scoring criteria have been reported in Merritt and Lies (1987) to enjoy the psychometric properties of reliability and validity.

Materials

Stories

The stories extracted from Stories Alive (Bilbrough, 2016) and published online by British Council were used in both classes. Stories were modified to become similar in number of story

episodes. The readability of stories was analyzed using Dale- Chall formula in order to control the condition for the sake of the study. Readability scores were 4.85 and 4.96 for the stories used in pre-test and post-test, respectively which indicated that texts were easily readable by average 4th grade. All stories had six to eight pictures and six to eight main events as story summaries. Stories were all culturally known to the participants as they were normally narrated in the Middle East countries like Juha and his Donkey. Except for the stories mentioned in pre- test and post-test, four more stories were picked up to work during the study (refer to Appendix C for the names of stories). Stories had well- developed story structures meaning three episodes containing characters, settings, initiating events, internal response, attempt and consequence (see Appendix B).

Procedure

After participant separation and research group formation as mentioned above, a pretest of oral proficiency and pretest of storytelling skill were administered and rated according to the rubric mentioned in the previous section. The treatment sections were held as detailed in subsections in each research group. Oral proficiency test and storytelling skills test were once again conducted as posttest to trace the effect of treatment. The treatments were delivered by the same researcher in both groups to reduce the teacher effect. The duration of the treatment sessions for both groups was the same (12 sessions excluding the sessions allocated to pre-test and post-test) to neutralize the time on task effect. The format of the lessons was identical for both groups. All in all, six stories were covered in sessions two, four, six, eight, ten and twelve while students were instructed to work on story grammar components in sessions five, seven, nine, eleven and thirteen. The story lessons started with lead- in activities, followed by main reading activities and ended with post reading. However, critical discussions and activities were incorporated into CP lessons as well. Subsequently, story grammar component lessons began with an introduction to a story grammar component and ended with related practices. See Appendix C for the number of the sessions, the name of the stories, grammar story elements and a short description of each session in both classes. In the first session, the researcher and participants primarily focused on getting to know each other and familiarized themselves with the content of the course. In addition, sessions three and fourteen were allocated to pre- tests and post-tests, respectively.

CP-Based Teaching Classroom

As Akbari (2008) claims, CP is not a method with a well-defined procedure. Therefore, CP-based techniques and activities proposed by Abednia (2015), Aghcelu (2013) and Akbari (2008) were extracted from literature to apply the practical implications of CP in L2 classroom. **I**n the first session, the instructor tried to break the ice with students, got to know them and talked about the benefits of storytelling. She tried to learn about students' beliefs, attitudes, social stands and problems as much as possible by opening up to them and questioning her own life-style as she was introducing herself to them. Before she precedes more, she informed students that she was going to make two mistakes deliberately by the end of the session and she would like the students to find them. Aghchelu (2013) mentioned that finding the teacher's error encourages the students to question and later checks what is being taught by the teacher in class and thus students do not trust the teacher as the only source of knowledge. They were also informed that they could resort to their L1 whenever they need to and took part in class discussions and other activities (Akbari, 2008).

Abednia (2015) once mentioned that one of the tenets of CP lessons is the active engagement of learners in choosing the texts they want to study. As an alternative, he also suggests that the teacher chooses the texts. In this study, a combination of both was used. One way to apply CP in L2 classroom is basing the lessons on students' local culture (Akbari, 2008). Thus, the teacher

chose a collection of short stories, named Stories Alive (Bilbrough, 2016) which were well-known to the Middle East like "the tortoise and the hare" and then introduced the stories to the students. Among ten stories, they chose six of them to work on. Consequently, learners' active participation in selecting the stories they wanted to read was assured.

During the second session, both classes worked on a story from Stories Alive. Students in CP class had critical discussions, approached the text in a questioning manner and analyzed the text of the story critically to see who benefits it the most. They tried to empathize with the characters and found solutions to their problems (Abednia & Izadinia, 2013). They also were asked if they had similar problems in their own life and were encouraged to come up with real solutions (Akbari, 2008). In the third session, students were given a set of photos of the story of the tortoise and the hare, which they had worked on previous session. Each student was asked to retell the story and their stories were recorded. After telling stories, each individual had two to three minutes to rest and get prepared for the oral proficiency test. From the third session on students were introduced to the elements of storytelling gradually.

Another activity found in the literature was assigning students for and against (Aghchelu, 2013). When students were working on the story of the Nasredin and the Dinner Party, students were asked to state their ideas on what they thought about the moral of the story. After students communicated their ideas in class, they were asked to choose one moral to discuss. Eventually, each student's idea was asked individually. Those who agreed with the selected message formed one group and those who disagreed, formed another group. Then, students were assigned to be for and against meaning that those who agreed with the message were assigned against and those who disagreed were assigned for. They had to work in groups and come up with reasons to justify the ideas assigned to them. This activity helped students to be open-minded and see the world from other people's perspectives and see if they can change their attitude and look at the world from a different perspective.

The next activity used was to pin the blame! (Aghchelu, 2013). In another session when they were working on "the Jackal and the Crow", students were requested to work in group and pin the blame on one of the main characters. Since there were two main characters in the story, some took the view that the jackal was guilty and the rest acknowledged that we should put the blame on the crow. Then, students from opposing groups were paired up and invited to defend their ideas. They were also encouraged to judge the main two characters and decided whose right was to eat the piece of chicken and have a debate with their partners and support their ideas by depicting real examples.

"Juha and his Donkey" was another story which prepared many opportunities for critical discussions. After reading the story and working on students' comprehension, they wrote their reflections on a piece of paper regarding whose character's words should be taken into account. Most students agreed that it was a wrong idea to listen to what people say including even the grandmother. Then the teacher encouraged the students to think about the cliché that children should always follow what older people tell them. The activity discussed above were mentioned in (Aghchelu, 2013).

Asking students to write reflections leads them to have critical evaluation. Students commented on each other's reflections without the teacher getting involved. This is good for creating an ambience in which the teacher is not the class hierarchy and students learn to refer to each other as resources and they are trained to be critical. (Osterfelt, 2011) Students were also required to write journals while using L1 was not condemn in CP class (Miri, Alibakhshi, & Mostafaei-Alaei, 2017).

Task-Based Teaching Classroom.

Task-Based lessons were planned based on the three main stages suggested by (Willis & Willis, 2007) which are pre-task, cycle task and post- task. The very first session was spent mostly on introduction and getting to know the students. The researcher first introduced herself and then led the students to familiarize with each other through a mingling activity and then, the purpose of the course and the chosen stories were introduced to the learners. It is also worth mentioning that while students in CP-based teaching class had their own say in choosing the stories, in TBLT group students' voices were not heard and the very same stories chosen in CP class were used in the TBLT class, as well.

The next session in the pre-task stage, the teacher explained that learners should work together to brainstorm the names of animals and write an adjective, which describes each animal next to it. Then, students read their lists out loud to the class while the teacher wrote them on the board. In the next stage, which was cycle-task, the teacher explained that they were going to learn a new story about a lion and a mouse. Sets of pictures depicting different scenes of the story were given to each group. Students were asked to work together again, described what they see in pictures, put them in order and guess the story. Later, students were provided with sets of written strips and were asked to match the strips with the pictures. Finally, from each group, one person volunteered to tell the story. At this time, students' interest and curiosity were aroused and they were ready to listen to the teacher telling the original story. As the teacher told the story, students were required to check if they had ordered the pictures correctly or not and figured out how different their stories were compared to the original ones. Subsequently, students worked together and answered a set of comprehension questions related to the story. The answers were checked in class and the group with most correct answers was chosen. As the post-task activity students drew a story line to summarize the main events and then practiced retelling the story in their groups and got ready to act it out.

The procedure in the other sessions was almost the same in terms of working on stories, however, since there was no need to spend time on having critical discussions, more time was at teacher and students' disposal to get involved in other exciting activities such as acting out the stories, drawing story maps, creating their own ending, answering comprehension questions, watching related clips and playing vocabulary games. Students worked on stories using matching tasks, predicting tasks, jigsaw reading, problem solving, opinion gaps, reasoning gaps etc. Unlike CP-based classroom, in TBLT class students were vehemently discouraged to use L1 and only judicious use of L1 was accepted. There was no sign of asking questions which encouraged students to judge and evaluate stories. Personalized questions were asked but not in a way that students had to find solutions to their real life problems.

Data Collection and Analysis

Pre-test and post- test were administered after one session working on the story of the tortoise and the hare in pre-test and the jackal and the crow in post-test (see appendix D). Each participant entered the room individually for telling their stories. Six pictures including the main events were given to them. They were first asked to put the pictures in order and when they felt ready, they started telling their stories. McKay (2006) believes that teachers should scaffold children when they are engaged in an oral assessment. Therefore, to assist participants when they were telling their stories standardized prompts were provided like "what happened next?" and "what happened in the end?" (Ma et al., 2017; Morrow, 1986). Students were allowed to ask for the words they did not remember or did not know.

Participants were asked to retell the stories for which each person used the same six pictures as story motivators. Different sets of figures were used in the pre-test and the post-test. However, in both pre- and post-tests, the figures included two animals as the main characters and some trees. Students could use all or only some of the pictures (see Appendix E). Their stories were taped, transcribed, and analyzed for inclusion of story elements.

Having told the story, each learner was given a two-minute recess. Immediately after the break, the test taker was prepared for the oral proficiency test. Learners were assessed through Cambridge Young Learners English Test (2014), which consisted of four parts. In part one, learners were asked to talk about the differences between two pictures. In second part, learners told a story based on the pictures given to them. Next, the learners had to find the pictures, which were different from sets of four pictures and explained the reason why he thought the picture was different. Finally, the learners had to answer open-ended questions. The same procedure was followed in the post-test session. The only difference was the story, its pictures and the oral proficiency test. As their post-test, students had to tell the story of the jackal and the crow one session after working on it.

Having calculated inter-rater and intra-rater reliability on oral proficiency tests and episode analysis for recorded stories told by the participants, the researchers conducted series of an independent sample t-test to locate significant differences between groups.

4. RESULTS

The present study was concerned with comparing the effects of Critical Pedagogy Based Teaching and Task-Based Language Teaching. This involved having two groups of language learners and comparing the results of their pretest/posttest research design to investigate any significant effects on participants' storytelling skills and oral proficiency before and after 12 sessions of treatment in each research group.

In order to estimate intra- rater and inter- rater reliability intraclass correlation coefficient was calculated. The intra- rater repeatability resulted from intraclass correlation was excellent for oral proficiency in pre-test and post-test being 0.994 and 0.988, respectively. Furthermore, having used the same model to calculate the intra-rater repeatability for story grammar components, intraclass coefficient was 0.993 in pre-test and 0.996 in post-test.

Descriptive statistics

Table 1 shows the basic summary statistics with the mean of 13.46 and standard deviation of 0.73 regarding the age and number of participants. The skewness for none of the variables is 0 which means the distribution is not perfectly normal which a common occurrence in social studies is. Given that our findings are based on a limited number of participants and it is a limitation of this experimental study, the results should be treated with considerable caution. However, since t-test is not very sensitive to violation of some assumptions including normality (Pallant, 2013 & Field, 2013), an independent-sample t-test was conducted to compare students' oral proficiency for CP and TBLT class after taking part in pre- test.

Table 1: Descriptive of participants in both groups for pre- tests and post- tests

	N	Mean	Std. Deviation
	Statistic	Statistic	Statistic
O1	30	11.90	3.25
O2	30	12.93	2.31
SG1	30	10.80	3.48
SG2	30	13.13	3.77
Age	30	13.46	.73
Valid N (Listwise)	30		

Table 2: Group Statistics of CP and TBLT groups in Oral Proficiency Post- test

Method		N	Mean	Std. Deviation	Std. Error Mean
	CP	15	12.60	2.35	0.60
Oral2	TBLT	15	13.26	2.31	0.59

TBLT group (12.53) is more than the mean score of CP group (11.26). In addition, standard deviation for the CP and TBLT groups are, respectively SD = 3.45 and SD = 3.02. The Sig. value for Levene's test (0.141) shows that equal variances should be assumed. Thus the Sig. (2- tailed value) bigger than 0.05 (p= 0.294) means there is no significant difference in the mean scores of two groups in oral proficiency pre-tests. Moreover, the Table shows that the magnitude of the differences in the means (mean difference= - 1.26, 95% CI:3.69 to1.15) is small (0.039).

Findings on oral proficiency in post- test.

Table 2 demonstrates the participants' mean scores and standard deviations for CP and TBLT groups in post-test. As Table 2 reveals, the figure for the mean score of TBLT group is more than that of CP group. While there is no significant difference in the standard deviation of both groups (SDCP = 2.35, SDTBLT = 2.31).

The mean scores of both groups have increased; however, as Table 3 depicts, the rise is not significant. The Sig. value for Levene's test is 0.635 which means we should refer to Sig. (2- tailed) for equal variances. This value (0.441) indicates that no group performed significantly better than the other group in oral proficiency post-test. In addition, Eta squared is 0.021 which means the effect size is small. It is worth mentioning that T-test effect size is not provided by IBM SPSS in the output. However, by knowing t-value and the number of participants in each group, it can be manually calculated using the following formula:

Eta squared$= (t^2 \div (N1 + N2 - 2)) = ((-0.78)^2 \div (30 + 30 - 2)) = 0.021$

Table 3: Independent Samples Test of CP and TBLT groups in Oral Proficiency Post - test

		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
				Levene's Test				
	Equal variances assumed			- 0.78	28	0.44	-0.66	0.85
		0.23	.63					
Oral2	Equal variances not assumed			- 0.78	27.99	0.44	-0.66	0.85

Findings on Story Grammar Components

Findings on Story Grammar Components in pre- test.

The second research question aimed to investigate the effects of Critical Pedagogy and Task-Based Language Teaching on Story setting, initiating events, internal response, attempt, consequence (see Appendixes A and B). The independent sample t-test performed on the data obtained from Story Grammar pre-tests displayed the group statistics, namely the number of participants, mean score and standard deviation for both CP and TBLT groups (Table 4).

As it is shown in Table 4, the TBLT group has a higher performance, (M = 11.86, SD = 2.89) followed closely by that of CP group (M = 9.73, SD = 3.79). The second part of the result section is related to independent samples test investigating the effects and it shows variances are equal and to check any significant differences between the results of both groups, we should use the first Sig. (2- tailed) related to the equal variances. Since the Sig. (2- tailed) is larger than 0.05, the difference between two groups is not noteworthy. Furthermore, there is moderation in the figure for Eta squared (0.096).

Table 4: Group Statistics of CP and TBLT groups in Grammar Components Pre- test

	Method	N	Mean	Std. Deviation
GC1	CP	15	9.73	2.89
	TBLT	15	11.86	3.79

Table 5: Independent Samples Test of CP and TBLT groups in Story Grammar Components Pre- Test

		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
				Levene's Test				
	Equal variances assumed			- 1.73	28	.094	-2.13	1.23
		0.18	.0.67					
GC1	Equal variances not assumed			- 1.73	26.14	.095	-2.13	1.23

Table 6: Group statistics of CP and TBLT groups in Story Grammar Components Post- Test

	METHOD	N	Mean	Std. Deviation	Std. Error Mean
	CP	15	12.33	3.79	.97
T2	TBLT	15	13.93	3.71	.95

Table 7: Independent Sample Test of CP and TBLT groups in Story Grammar Components Post-Test

Levene's Test			t-test for Equality of Means				
	F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Equal variances assumed T2	1.88	.181	- 1.16	28	.253	-1.60	1.37
Equal variances not assumed			-1.16	27.98	.253	-1.60	1.37

Findings on Story Grammar Components in post-test.

Table 6 contains the information about group statistics in post-tests aimed to investigate CP and TBLT impacts on story grammar components. As it can be seen, there is an increase in CP mean score from 9.73 to 12.13. Similarly, TBLT mean score has approximately risen by 2. This is true while there is no significant difference between the standard deviation of both groups (3.79 for CP group and 3.71 for TBLT).

The Sig. value (0.18) delineated in Table 7 suggests that the Sig. (2- tailed) allocated to equal variances should be used. According to Sig. (2- tailed) which is 0.253 and obviously higher than 0.05, the difference between CP and TBLT groups in terms of inclusion of story grammar components in students' stories is not noticeable. In addition, the magnitude of the difference is small (eta squared = 0.045).

Results of Complete Episodes

Results of Complete Episodes in Pre-test.

The comparison of the data displayed in Table 8 shows that the mean score of TBLT group (1.2) is larger than the mean score of CP group (0.8); however, this difference is insignificant.

Table 9 displays that considering the figures for p value and Sig. (2- tailed) which are 0.779 and 0.214, respectively, there is no significant difference in the number of complete episodes between the two groups in the pre-test phase. The magnitude of the differences in the means (0.056) indicates a small effect size, as well.

Table 8: Group Statistics of CP and TBLT group in Complete Episodes Pre -Test

	METHOD	N	Mean	Std. Deviation	Std. Error Mean
CE1	CP	15	.80	.77	.20
	TBLT	15	1.20	.94	.24

Table 9: Independent Samples Test of CP and TBLT groups in Complete Episodes Pre- Test

		Levene's Test		t-test for Equality of Means				
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
CE1	Equal variances assumed	.08 0	.77	-1.27	28	.214	-.40	.314
	Equal variances not assumed			-1.27	27.00	.215	-.40	.314

Findings on Complete Episodes in Post- test.

The mean scores for the number of complete episodes are shown in Table 10 as follows CP (M= 1.13, SD = 0.47) and TBLT (M = 1.4, SD=0.73) in post- test.

Table 10: Group Statistics of CP and TBLT group in Complete Episodes, Post-Test

METHOD	N	Mean	Std. Deviation	Std. Error Mean
CP	15	1.13	.74	.19
CE2				
TBLT	15	1.40	.73	.19

The findings of the comparison between the mean scores (P value= 0.663, Sig. (2- tailed) = 0.332), however, fails to show any significant difference between the performance of the participants in post-test in terms of the number of completed episodes (Table 11). Eta squared=0.03 illustrates a small effect.

The differences of the study, therefore, are insignificant between Tasked-Based Language Teaching and Critical Pedagogy based teaching. The findings of the study must be treated with caution, due to the sample size of the study and reliability. Figure 1 shows the results acquired during the whole study.

Table 11: Independent Samples Test of CP and TBLT groups in Complete Episodes Post – Test

		Levene's Test for		t-test for Equality of Means				
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
CE2	Equal variances assumed	.19	.66	-.98	28	.33	-.26	.27
	Equal variances not assumed			-.98	27.99	.33	-.26	.27
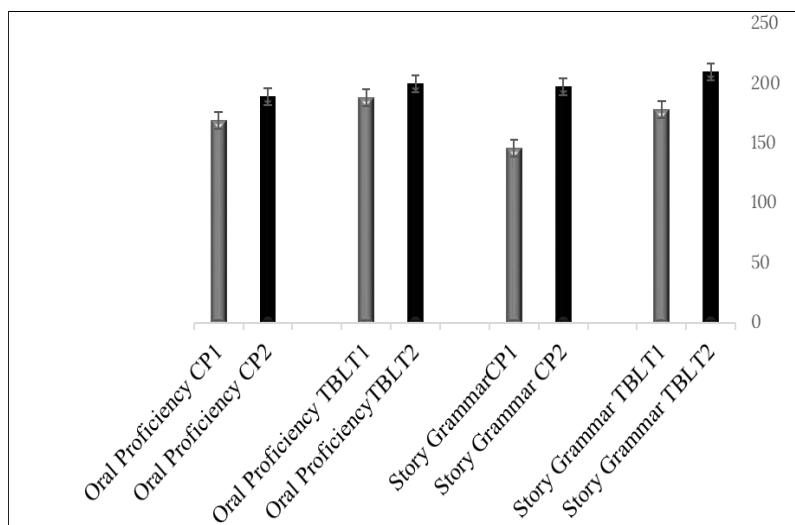
Figure 1: Comparison of the results acquired from pre- tests and post- tests for treatment and controlled groups

5. DISCUSSION

This study aimed at investigating the potential of two instruction methods of CP and TBLT on the development of student oral proficiency and storytelling skills. The results indicated that there were not any significant differences between the potential of the two on the development of students' oral proficiency and storytelling skills meaning that both methods were somehow effective to the same extent.

The results of the study are supported by several similar studies as far as the efficacy of task-based and critical pedagogy with respect to oral proficiency was concerned. For example, Murad and Smadi (2009) indicated that TBLT was efficient in improving Palestinian secondary EFL students. Likewise, Albino (2017) confirmed the positive effect of TBLT on speaking fluency of EFL learners in PUNIV- Cazenga. SimilarlyAshraf Ganjouee et al. (2018)'s, study of the efficacy of TBLT on speaking skills of Iranian EFL learners was in line with the present research results. Aligned with the above-mentioned studies, the results of the present study also corroborate a number of other influential studies as far as Critical Pedagogy is concerned. Riasati and Mollaei (2012)'s analysis of the critical pedagogy and EFL learning approved its positive relation. Similarly, Noroozisiam and Soozandehfar (2011)'s analysis of the efficacy of CP on EFL and ESL speaking skill showed achievement in Iranian language learners. Likewise, Mohammadi et al.'s (2014) study results were in line with the present research since both approved the positive effect of CP on oral proficiency.

The results of the study were also consolidated with other research findings as far as the efficacy of TBLT and critical pedagogy was concerned with respect to storytelling skills. Ahmed and Bidin (2016) approved the effect of TBLT on storytelling skills and narrations of Malaysian EFL learners. However, their study focused on storytelling skills in written modality. Likewise, Rodríguez Buerba's (2019) study results indicated that that the practice of task-based storytelling has positive effects on EFL elementary learners' speaking anxiety and higher speaking performance achievement. As far as the efficacy of CP on storytelling was concerned, the study results corroborated the results of the study by Enciso (2011) in which the positive effect was

confirmed. In a comparative sense, none of the instruction methods of CP and TBLT was in privilege over the other and the results of the present study was not in favor of one over the other.

The results were also supported and backed up with socio constructs approaches towards learning. As Vygotsky (1987) stated it is through learner engagement via interaction and negotiation of meaning that learners can co- regulate their learning within their zone of proximal development and scaffolding can change inter-subjectivity to intra-subjectivity (self-regulation). Co-regulation helps students to internalize students' self-regulation potential through interaction with more competent peers within students' zone of proximal development. Interaction can create moments of contingency in instruction for the purpose of the regulation of learning processes (Black & Wiliam, 2009). Regulated learning theories and models have been formed to include situated perspectives of learning; perspectives that highlight highly interactive and dynamic learning situations where shared knowledge construction and collaboration emerge. Self-regulated learning became a foundation in exploration of more social forms of regulation such as co-regulation and shared regulation (Hadwin et al. 2018).

6. CONCLUSION AND IMPLICATIONS

The present research reveals the fact that the Critical Pedagogy as a recent change and Task-Based Pedagogy as a common methodology did not differ greatly. Although there are many recommendations regarding use of the Critical Pedagogy in ELT, the result of this study did not depict any outstanding benefits to its use and in cases where there is not enough time to see the long-term effect of CP, perhaps neither Task- Based Language Teaching nor Critical Pedagogy Based Teaching has the upper hand regarding students' improvement of storytelling and oral proficiency. This, in fact, highlights the vital need for more practical and quantitative studies in ELT so that we can assure efficiency in our language classes.

The analysis of the data also reflects the fact that a longer period of time and a larger sample group is required for more in-depth studies. Since the treatment was short, we are also of the opinion that the CP course could not meet its aim which was transformation. This claim is in line with previous studies like (Ro, 2015). In order to see more long-term impacts, long-running studies are required (Korthagen, 2004). If the course was prolonged, the instructor and students were likely to find better opportunities to practice the cornerstone of CP in students' real lives.

One of the tenets of CP was validating students' first language and not prohibiting them from using L1. However, there was considerable concern for the detrimental influence of using L1 over students' fluency and developing the habit of using their mother's tongue and excessive code switching in class, students' fluency improved and there was no outstanding difference noticed.

Hinged on what research showed, a glory gap was felt for more investigations owing to participants' age and language level. It was pointed out that doing more practical studies was vital on the subject of implementing Critical Pedagogy and developing students' critical thinking in classroom (Akbari, 2008). Both the level of language and their age were important factors which needed further research.

References

Abednia, A., & Izadinia, M. (2013). Critical pedagogy in ELT classroom: Exploring contributions of critical literacy to learners' critical consciousness. *Language Awareness, 22*(4), 338-352.

Abednia, A. (2015). Practicing critical literacy in second language reading. *The International Journal of Critical Pedagogy, 6*(2).

Adiantika, H. N., & Purnomo, H. (2018). The implementation of task-based instruction in EFL teaching speaking skill. *Indonesian EFL Journal, 4*(2), 12-22.

Aghchelu, S. (2013). Teachers' familiarity with critical pedagogy and their critical thinking abilities. (Unpublished master's thesis). Allameh Tabataba'i, Iran.

Ahmed, R. Z., & Bidin, S. J. B. (2016). The effect of task-based language teaching on writing skills of EFL learners in Malaysia. *Open Journal of Modern Linguistics, 6*(3), 207-217. https://doi.org/10.4236/ojml.2016.63022

Akbari, R. (2008). Transforming lives: Introducing critical pedagogy into ELT classrooms. *ELT Journal, 62*(3), 276-283. https://doi.org/10.1093/elt/ccn025

Albino, G. (2017). Improving speaking fluency in a task-based language teaching approach: The case of EFL learners at PUNIV-Cazenga. *SAGE Open, 7*(2), Article 2158244017691077. https://doi.org/10.1177/2158244017691077

Ansarin, A. A., & Mohamadi, Z. (2013a). Language engagement at the level of syntax: Assessing metatalk and task types in SLA. *International Journal of Applied Linguistics and English Literature, 2*(4), 142-154. https://doi.org/10.7575/aiac.ijalel.v.2n.4p.142

Ansarin, A. A., & Mohamadi, Z. (2013b). Language engagement in task-based interaction: Focus on intonation. *The Iranian EFL Journal, 12*(2), 152-173.

Ashraf Ganjouee, A., Ghonsooly, B., & Hosseini Fatemi, A. (2018). The impact of task-based instruction on the enhancement of Iranian intermediate EFL learners' speaking skill and emotional intelligence. *Applied Research on English Language, 7*(2), 195-214.

Assalahi, H. M. (2013). Why is the grammar-translation method still alive in the Arab world? Teachers' beliefs and its implications for EFL teacher education. *Theory & Practice in Language Studies, 3*(4), 589-599.

Atai, M. R., & Moradi, H. (2016). Critical pedagogy in the context of Iran: Exploring English teachers' perceptions. *Applied Research on English Language, 5*(2), 121-144.

Bain, J. (2010). Integrating student voice: Assessment for empowerment. *Practitioner Research in Higher Education, 4*(1), 14-29.

Bilbrough, N. (2016). *Stories Alive* (pp. 109). Retrieved from https://www.britishcouncil.ps/sites/default/files/storiesalive_-final_palestine_version.pdf

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Boegeman, W. S. (2013). Teaching for transformation: The praxis of critical pedagogy in social studies education. (Unpublished doctoral dissertation). University of Minnesota.

Britton, E. R., & Leonard, R. L. (2020). The social justice potential of critical reflection and critical language awareness pedagogies for L2 writers. *Journal of Second Language Writing, 50*, Article 100776. https://doi.org/10.1016/j.jslw.2020.100776

Brown, H. D. (2014). *Principles of language learning and teaching* (6th ed.). Pearson Education.

Cambridge. (2001). *Quick Placement test*. University of Cambridge: Oxford University Press and Umiversity of Cambridge Local Examination Syndicate.

Cambridge. (2013). *Young Learner English Tests*. Retrieved from https://www.cambridgeenglish.org/?153310-movers-sample-papers-volume-2.pdf

Cambridge. (2014). *Young Learners English Test*. Retrieved from https://www.cambridgeenglish.org/?153310-movers-sample-papers-volume-2.pdf

Cambridge. (2018). *Handbook for Teachers*. Retrieved from https://www.cambridgeenglish.org/de/Images/357180-starters-movers-and-flyers-handbook-for-teachers-2018.pdf

Crookes, G. (2012). Critical pedagogy in language teaching. In *The Encyclopedia of Applied Linguistics*. Wiley. https://doi.org/10.1002/9781405198431.wbeal0265

Cutrone, P., & Beh, S. (2018). Investigating the effects of task-based language teaching on Japanese EFL learners' willingness to communicate. *Journal of Asia TEFL, 15*(3), 22-38. https://doi.org/10.18823/asiatefl.2018.15.3.3.22.

Ellis, G., & Brewster, J. (2014). *Tell it again!: The storytelling handbook for primary English language teachers*. British Council.

Ellis, R., Skehan, P., Shintani, N., & Lambert, C. (2019). *Task-based language teaching: Theory and practice*. Cambridge University Press.

Enciso, P. (2011). Storytelling in critical literacy pedagogy: Removing the walls between immigrant and non-immigrant youth. *English Teaching: Practice and Critique, 10*(1), 21-40.

Field, A. (2013). *Discovering statistics using SPSS* (3rd ed.). SAGE Publications.

Ghahremani-Ghajar, S., & Mirhosseini, S. A. (2005). English class or speaking about everything class? Dialogue journal writing as a critical EFL literacy practice in an Iranian high school. *Language, Culture and Curriculum, 18*(3), 286-299. https://doi.org/10.1080/07908310508668749

Gruenewald, D. A. (2003). The best of both worlds: A critical pedagogy of place. *Educational Researcher, 32*(4), 3-12.

Hadwin, A., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Educational Psychology Handbook Series. Handbook of Self-Regulation of Learning and Performance* (pp. 83–106). Routledge.

Hasan, A. A. A. (2014). The effect of using task-based learning in teaching English on the oral performance of the secondary school students. *International Interdisciplinary Journal of Education, 3*(2), 250-264.

Hawkins, M., & Norton, B. (2009). Critical language teacher education. In *Cambridge Guide to Second Language Teacher Education* (pp. 30-39). Cambridge University Press.

Korthagen, F. A. (2004). In search of the essence of a good teacher: Towards a more holistic approach in teacher education. *Teaching and Teacher Education, 20*(1), 77-97. https://doi.org/10.1016/j.tate.2003.10.002

Kumaravadivelu, B. (2006). TESOL methods: Changing tracks, challenging trends. *TESOL Quarterly, 40*(1), 59-81. https://doi.org/10.2307/40264511

Larsen-Freeman, D., & Anderson, M. (2013). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.

Lin, Z. (2009). Task-based approach in foreign language teaching in China: A seminar paper research presented to the graduate faculty, University of Wisconsin-Platteville. Retrieved October 15, 2014.

Ma, S., Anderson, R. C., Lin, T.-J., Zhang, J., Morris, J. A., Nguyen-Jahiel, K., & Sun, J. (2017). Instructional influences on English language learners' storytelling. *Learning and Instruction, 49*, 64-80. https://doi.org/10.1016/j.learninstruc.2017.12.008

 ALIKAHI, F., & KIANY, G. R.

Mahmoodarabi, M., & Khodabakhsh, M. R. (2015). Critical pedagogy: EFL teachers' views, experience and academic degrees. *English Language Teaching, 8*(6), 100-113. https://doi.org/10.5539/elt.v8n6p100

McKay, P. (2006). *Assessing young language learners*. Cambridge University Press.

McLaren, P. (1992). Critical literacy and postcolonial praxis: A Freirian perspective. *College Literature, 19*(3/1), 7-27.

Merritt, D. D., & Liles, B. Z. (1987). Story grammar ability in children with and without language disorder: Story generation, story retelling, and story comprehension. *Journal of Speech, Language, and Hearing Research, 30*(4), 539-552. https://doi.org/10.1044/jshr.3004.539.

Miri, M., Alibakhshi, G., & Mostafaei-Alaei, M. (2017). Reshaping teacher cognition about L1 use through critical ELT teacher education. *Critical Inquiry in Language Studies, 14*(1), 58-98. https://doi.org/10.1080/15427587.2016.1141803

Mohamadi, Z. (2015). Negotiation of meaning in required and optional information exchange tasks: Discourse issues. *International Journal of Applied Linguistics and English Literature, 4*(1), 95-105. https://doi.org/10.7575/aiac.ijalel.v.4n.1p.95

Mohamadi, Z. (2017a). Task engagement: A potential criterion for quality assessment of language learning tasks. *Asian-Pacific Journal of Second and Foreign Language Education, 2*(1), Article 3. https://doi.org/10.1186/s40862-017-0024-z

Mohamadi, Z. (2017b). Task engagement: A potential criterion for quality assessment of language learning tasks. *Asian-Pacific Journal of Second and Foreign Language Education, 2*(1), Article 1. https://doi.org/10.1186/s40862-017-0035-8

Mohamadi, Z., & Rahimpour, M. (2018). Task types and discursive features: Mediating role of meta-talk in focus. *Iranian Journal of Language Teaching Research, 6*(1), 17-40.

Mohammadi, N., Motallebzadeh, K., & Ashraf, H. (2014). Critical pedagogy: A key factor for improvement of Iranian EFL learners' self-regulation and writing ability. *International Journal of Applied Linguistics and English Literature, 3*(4), 46-54. https://doi.org/10.7575/aiac.ijalel.v.3n.4p.46

Morrow, L. M. (1986). Effects of structural guidance in story retelling on children's dictation of original stories. *Journal of Reading Behavior, 18*(2), 135-152. https://doi.org/10.1080/10862968609547527

Murad, T. M., & Smadi, O. (2009). The effect of task-based language teaching on developing speaking skills among the Palestinian secondary EFL students in Israel and their attitudes towards English. *The Asian EFL Journal*.

Naiditch, F. (2016). Critical pedagogy and the teaching of reading for social action. In *Developing Critical Thinking: From Theory to Classroom Practice* (p. 87).

Noroozisiam, E., & Soozandehfar, S. M. A. (2011). Teaching English through critical pedagogy: Problems and attitudes. *Theory & Practice in Language Studies, 1*(9), 1240-1244. https://doi.org/10.4304/tpls.1.9.1240-1244

Oradee, T. (2012). Developing speaking skills using three communicative activities (discussion, problem-solving, and role-playing). *International Journal of Social Science and Humanity, 2*(6), 533-535. https://doi.org/10.7763/IJSSH.2012.V2.164

Osterfelt, C. (2011). Critical dialogues: An action research project and a critical pedagogy professional development group for public school teachers in Peoples, AZ. (MA thesis). Prescott College.

Pallant, J. (2013). *SPSS survival manual*. McGraw-Hill Education (UK).

Riasati, M. J., & Mollaei, F. (2012). Critical pedagogy and language learning. *International Journal of Humanities and Social Science, 2*(21), 223-229.

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

Ro, E. (2015). Examining critical pedagogy with drama in an IEP context. *Journal of Language Teaching and Research, 7*(1), 1-8. https://doi.org/10.17507/jltr.0601.01

Rodríguez Buerba, P. (2019). A strategy to reduce the negative effects of anxiety in EFL elementary learners' speaking. Universidad Veracruzana.

Sadeghi, S. (2008). Critical pedagogy in an EFL teaching context: An ignis fatuus or an alternative approach. *Journal for Critical Education Policy Studies, 6*(1), 277-295.

Shin, J. K., & Crandall, J. (2014). *Teaching young learners English: From theory to practice*. National Geographic Learning.

Smith Byron, A. (2011). Storytelling as loving praxis in critical peace education: A grounded theory study of postsecondary social justice educators. Portland State University.

Stein, N. L., & Glenn, C. G. (1975). An analysis of story comprehension in elementary school children: A test of a schema. *ERIC Document Reproduction Service.*

Takala, A. (2016). Grammar teaching methods in EFL lessons: Factors to consider when making instructional decisions. (Master's thesis). University of Jyväskylä, Finland.

Van de Guchte, M., Rijlaarsdam, G., Braaksma, M., & Bimmel, P. (2019). Focus on language versus content in the pre-task: Effects of guided peer-video model observations on task performance. *Language Teaching Research, 23*(3), 310-329. https://doi.org/10.1177/1362168817734297

Vogel, D., Rudolf, M., & Scherbaum, S. (2020). The impact of the verbal instruction and task characteristics on effect-based action control. *Cognitive Processing, 21*(4), 607-621. https://doi.org/10.1007/s10339-020-00974-6

Vygotsky, L. (1987). Zone of proximal development. In *Mind in society: The development of higher psychological processes* (pp. 157-161). Harvard University Press.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.

Wright, A. (1995). *Storytelling with children*. Oxford University Press.

Zenouzagh, Z. M. (2020). Syntactic complexity in individual, collaborative, and E-collaborative EFL writing: Mediating role of writing modality, L1, and sustained development in focus. *Educational Technology Research and Development, 68*(6), 2939-2970. https://doi.org/10.1007/s11423-020-09774-w

Zhang, S., & Zhang, L. J. (2021). Effects of a xu-argument based iterative continuation task on an EFL learner's linguistic and affective development: Evidence from errors, self-initiated error corrections, and foreign language learning attitudes. *System*, Article 102481. https://doi.org/10.1016/j.system.2021.102481

Appendix A

Story component criteria

A statement is categorized as a SETTING if:

1. A major or minor character is introduced, (e.g., "Jim had been a truck driver for twenty years."), or
2. A location is described, or
3. Additional information is presented that conveys the habitual social context (e.g., "The boys were usually gone from home for only a short time."), physical context (e.g., "Life on the island was hard."), or

4. A character's habitual state is noted, the estate may not have been caused by any previous occurrence, and may not cause a subsequent event to happen, (e.g., "He was a very careful driver.")

Initiating Event Category

A statement is scored as an INITIATING EVENT if it begins a goal- based episode sequence in the story and causes the main character to respond. Initiating Event includes three types of information:

1. A character's action or an event (e.g., " The boys couldn't find fresh water of food.")
2. Natural occurrences, which are changes in the physical environment not caused by animate being (e.g., "One day it had been snowing for many hours.")

Internal events, including a character's internal perception of an external event (e.g, "One day they spotted a ship."), or, changes in the character's internal physiological state (e.g., " By noontime, it was getting harder and harder to breathe.")

Setting and Initiating Events are distinguished from each other in that the setting provides the context for the story and the Initiating Event always evokes an immediate response from the character.

Statements that are general events and do not lead to a goal are not scored, for example, "And then they drank water… an then they ate…"

Internal Response Category

A statement is characterized as an INTERNAL RESPONSE if it meets three criteria:

1. It describes the character's psychological state including emotions, goals, desires, intentions, or thoughts, for example, "He wanted to get home safely," or, "But they knew they could survive if they worked together," or, "They thought that the driver of the truck might be dead," AND,
2. It is casually related to an Initiating Event in the story, AND,
3. It leads to a plan sequence.

Attempt Category

A statement is categorized as an ATTEMPT if it represents a character's overt action toward resolving the situation or achieving a goal. For example, in Buried Alice, the main character attempts to get air into his truck by lighting a blowtorch and cutting a hole in the roof.

There needs to be a direct causal link or enablement relation between the Attempt and Either the Initiating Event or Internal Response that usually proceeds it, or a direct causal link or enablement relation between the Attempt and subsequent Direct Consequence.

Direct Consequence Category

A statement is categorized as a DIRECT CONSEQUENCE if it marks the direct attainment of the character's goal and is the result of one or more Attempt statements. A Direct Consequence usually leads to a character's reaction, but this may be unstated in the story.

Direct Consequence include three types of information:

1. Natural occurrences that influence the resolution of the story by facilitating or impeding attainment of the character's goal, (e.g., "One day the rainy season began.").
2. A character's action that results in either the attainment of a goal or a change in the sequence of events. For example, in shipwrecked, the children built a simple cabin to meet their goal of being sheltered from the rain.
3. End states, (e.g., "They were happy inside the cabin.")

Reaction Category

REACTIONS define how a character feels about the attainment or nonattainment of a goal (e.g., "Jim was relieved."), what the character thinks about it (e.g., "They knew they were lost."). or an action that is emotional (e.g., "They shouted their thanks.')

A reaction statement is causally linked to a Direct Consequence, which is usually the preceding statement. Occasionally, a Reaction precedes a Direct Consequence, but the causal connection between the statements need to be apparent. Reactions usually occur at the end of an episode, but they can also be inserted at the other points, for example, if a character pauses to reconsider a consequence and then proceeds.

Appendix B

Story retelling scoring procedure

Statements elicited during the story retelling task are included in the analysis if they meet criteria for a specific story category presented in Appendix D and if the following two conditions are met:

1. The statement has occurred in the original story. An exact replica is not required, but the retold statement needs to contain the same semantic content as the first version. For example, " He was not dead," conveys the same meaning as, "He was alive." Also, if details are omitted, e.g., numbers, specific times, etc., but the same story information is expressed, then the statement is scored.
2. The retold statement also needs to express the same story information, i.e., the same story component (e.g., Initiating Event), as the original version.

Five additional general scoring procedures are also routinely followed:

1. Only one statement is scored when a child uses two or more clauses to express information that had been presented in only one statement in the original story.
2. If a child uses one clause to express information conveyed in two separate statements from the original story, e.g., " So he pulled over and fell asleep," and two distinct story categories are expressed, in this case, an Attempt plus a Direct Consequence, the statement is scored as both categories.
3. When a statement is expanded upon later in the retold story, or self- corrected, only the expanded/ corrected version is scored.
4. Word finding errors are not penalized, e.g., "Bill" for "Jim," "bus" for truck," "blowerthing" for "blowtorch," "hook for "anchor," etc. (Graybeal, 1981)
5. Syntax errors, e.g., "And they dig," for "They started digging the snow," are also not penalized.

Statements are not included in the analysis if any one of the following conditions is noted:

1. A general comment or question unrelated to the story,
2. Repetition of a thought,
3. An unfinished statement that conveys an incomplete thought,
4. False start,
5. Formal endings,
6. Unclear statements in which the formation is not specific enough, irrelevant, or contradictory to the original story.
7. Extraneous information not presented in the original story, e.g., additional conflicts, plans, etc.,

8. Statements that convey only part of the information in the original story. For example, " and so they got up," is not the same as, " the boys woke up frightened," as the internal response of fear is not conveyed.

9. Statements conveying information that was assumed or implied in the original story, for example, "One day, Jim was riding along," is implied in the first story.

10. Statements presented in the wrong sequence such that a different intent and story category is expressed relative to the original story. Occasionally, a child expresses the sequence of events in an order different from the original story but consistent with the meaning of the story, for example, expressing the Initiating Event and then Setting, or a Reaction followed by a Direct Consequence end state. These statements are scored correct if the causal link is established. On other occasions, the sequence of events is wrong, the correct story category is expressed in the statements, but the story line is not logical. For example, a child who says, then they went on an island," (a Direct Consequence), and several statements later says, "But then they saw the island (an Initiating Event). The statements are not scored if this occurs because the relationship between the story parts is not appropriate.

Appendix c

The number of the sessions, type of lessons, covered content and short descriptions of CP and TBLT lessons

Sessions	Lesson Type	Covered Content	TBLT Lesson Description	CP Lesson Description
1	Introduction		Ice breaker Introduction to the content	Ice Breaker Introduction to the content, rules and expectations of the course Choosing the stories Students were asked to find the teacher's mistake during the course
2	Story lesson 1	The Tortoise and the Hare	Working on the first story based on TBLT	Working on the first story based on CP tenets Incorporating critical discussions into the lesson Analysing and questioning the text Coming up with solutions for real life problems Journal writing
3	Pre- test	Storytelling and oral proficiency test	Storytelling and oral proficiency tests were given to the learners	Students took the storytelling and oral proficiency tests
4	Story lesson 2	The Lion and the Mouse	The second story was covered based on TBLT	Working on the second story Critical discussions Journal writing
5	Grammar component lesson 1	Setting	Introducing the concept of setting Categorizing the related parts of the first and second stories into setting Making a list of animals and places for students' story settings Brainstorming adjectives to describe animals and places Coming up with information describing the habitual states of the characters or contexts	Introduction of the first story grammar component, setting, followed by several practices Students were challenged: Why did they come up with specific adjectives to describe the animals? for instance: Sly fox Students were encouraged to become aware of their biases and pre- established beliefs towards different groups of people in

The number of the sessions, type of lessons, covered content and short descriptions of CP and TBLT lessons

				society Journal writing
6	Story lesson 3	Juha and the Donkey	The third story was covered based on TBLT	Working on the third story Critical discussions Reflection and critical evaluation of story Challenging cliché Journal writing
7	Grammar component lesson 2	Initiating events	Introduction of the second story grammar component: Initiating events Categorizing the related parts of the first, second and third stories into initiating events	Introduction of the second story grammar component, initiating events, followed by several practices Students were challenged to think how different initiating events in the story may affect the readers'
			Extra practices Deciding on the initiating events of their own stories	judgement and the group who may benefit from the story Journal writing
8	Story lesson 4	The Boy who cried wolf	The fourth story was worked on based on TBLT approach	Working on the fourth story Pin the blame on the characters Assigning students into "For and Against" groups to analyse the story from different perspectives Journal writing
9	Grammar component lesson 3	Internal response	Introduction to the third story grammar component, internal response, followed by several practices.	Introduction to the third story grammar component, internal response, followed by several practices Categorizing the related sentences of the previous stories into internal response Assigning learner to "For and Against" groups to discuss the motives of the characters for their internal responses Journal writing
10	Story lesson 5	Nasreddin and the dinner party	Task Based stages namely pre-task, cycle- task and post- tasks were taken	Story was read and analysed Students gave their own reflections on the story Students challenged the cliché Journal writing
11	Grammar component lesson 4	Attempt	Introduction of the fourth story grammar component, attempt, followed by several practices Categorizing the related sentences of the stories into attempt	The story grammar attempt was introduced Students worked together on the previous stories to recognize the sentences which could be categorized as attempts Learners tried to see what differences would have been made based on different characters' actions Journal writing
12	Story lesson 6	The Jackal and the Crow	The teacher used TBLT based activities to help the learners read and understand the story	After reading the story and working on its meaning, students got engaged in " Pin the Blame!" activity Journal writing
13	Grammar component	Direct consequences	The last story grammar component was introduced and	Direct consequences as the last story grammar component were

The number of the sessions, type of lessons, covered content and short descriptions of CP and TBLT lessons

	lesson 5		practiced Students were asked to write a story and incorporate the story grammar components They were informed about the post- test in the following session	taught and practiced Students were instructed to reflect on the whole course Learners were requested to write a story integrating what they had learned so far Students were informed about the upcoming post- tests
14	Post- test	Storytelling and oral proficiency tests	Participants took part in the post- tests	Participants took part in the post- tests

Teachers' and Students' Attitude towards Code-Switching in Learning English in Iranian EFL Classes

Shahla Nazeri[1,*] (iD), Davoud Amini[2], and Farzad Salahshoor[2]

MA in TEFL, Azarbaijan Shahid Madani University, Tabriz, Iran
[2] Azarbaijan Shahid Madani University, Tabriz, Iran

Abstract
Attitudes toward L1 use in EFL classes have ranged from an absolute ban to an inherent aspect of effective teaching. Code-switching, which refers to particular instances of L1 use in relation to psychological and social demands of classroom communications, has been investigated in terms of its appropriateness with regard to various classroom conditions. The present study sought to explore the effectiveness of code-switching in EFL classes from both teachers' and learners' attitudes with a focus on the differential effects of a set of individual difference factors. The data were collected from 400 subjects, including 374 students and 26 teachers by means of questionnaires and an observation checklist. The results of the data analysis revealed that students had dominantly positive attitudes about different aspects of code-switching. Meanwhile, learners' age, gender, and social class were significant sources of differentiation in this regard. Apart from rather cynical attitudes of the teachers toward code-switching, more than half of them thought that students did not become fully dependent on code-switching for better understanding, and assumed that code-switching strengthened learners' English. The teachers' attitudes about code-switching did not differ due to the level of the class they were teaching. The findings provide further evidence for the effectiveness of code-switching in EFL classes.

1. INTRODUCTION

The development of English proficiency is an important goal for Iranian students. Besides studying English as a course at school, most of them decide to attend language institutes. This demand for learning English as a foreign language makes researchers think about different factors which can affect the process of teaching and learning. One of the controversial issues in this case is the use of first language in class. As stated in Bozorgian and Fallahpour (2020), "the issue of L1 use in L2/FL instruction has been permanently debated over the past decades." (p. 2). Although there have always been theories about minimizing L1 as much as possible (Turnbull, 2001),

* **Corresponding Author**: Shahla Nazeri, MA in TEFL, Department of English Azarbaijan Shahid Madani University, Tabriz, Iran. **Email**: nazeri.shahla74@gmail.com

nowadays researchers believe in L1 use as a facilitating tool in FL teaching and learning (Cook, 2001) which does not hinder the FL process (Miles, 2004).

The fundamental issue of code-switching draws investigators' attention since pure L1 use is not usually common in language classes. According to Lin (2013), the studies related to code-switching are usually conducted in two kinds of contexts: 1) L2 contexts like English as second language classes, 2) bilingual education classrooms. As a pedagogical instrument, code-switching is defined as the "alternation of two languages within a single discourse, sentence or constituent" (Poplack, 2000, p. 224). Although this alternation of languages in routine speech may disrupt and bewilder the interlocutors, it could be a constructive instrument (Ghafar Samar & Moradkhani, 2014) for both teachers and students to make most of the teaching-learning process. However, the use of code-switching is usually neglected in language classes regardless of positive impacts on students' learning. Thus, most of the time, the regulations held by institutes oblige the teachers to teach just in English and not to let the students talk in any other language. Accordingly, the teachers force their students to restrict themselves to English in talking and asking their questions during the class time. These methodological impositions on the teachers, and consequently on the learners, may have roots in the belief that learning a foreign language is best achieved when the L1 use is abstained in classes and teaching should be in TL exclusively (Bozorgian & Fallahpour, 2015).

Based on the previous research, the optimal amount of L1 use in the FL classes for the success of instruction process (Lin, 2013) is recommended. Consequently, code-switching is treated as a teaching strategy not a sign of the teachers' imperfection (Ahmad, 2009). Then, considering code-switching as a pedagogical instrument leads us to think about its trace in language classes more precisely. Apart from the benefits and defects of code-switching, which is going to be discussed in the next part, its impact on the acceleration of learning a foreign language (FL) relies, to a large extent, on the degree to which both teachers and learners are aware of the effectiveness of code-switching. Therefore, the researchers of the present study aim to investigate code-switching from EFL teachers' and learners' perspective. The following research questions were proposed:

1. What are Iranian EFL learners' attitudes about the effect of code-switching on their learning?
2. Do Iranian EFL learners with different age, gender and social class have different attitudes about the effect of code-switching on their learning?
3. What are Iranian EFL teachers' attitudes about the effect of code-switching on their students' learning?
4. Do Iranian EFL teachers teaching different levels of proficiency have different attitudes about the effect of code-switching on their students' learning?

More specifically, attempts were made to investigate EFL teachers and students' attitudes about this pedagogic element in language classes.

2. LITERATURE REVIEW

As Beatty-Martínez and Dussi (2017) have asserted "a unique feature of bilingual communication is that many bilinguals sometimes alternate between languages when speaking to other bilinguals" (p. 173). This alternation, which is a controversial issue in language classes is code-switching (Grosjean, 1982). According to Mirhasani and Jafarpour (2009), during the 1970s and 1980s, code-switching was viewed as a deficiency in language teaching and it was tried to be prevented in order to optimize teaching and learning by maximally using target language. However, code-switching, perforce, was accepted as an inescapable part of language classes, and since then "codeswitching in general has been a subject of great scholarly attention in recent decades" (Mirhasani & Jafarpour, 2009, p. 23).

Based on theoretical literature in bilingualism research, code-switching is a central issue in this area (Milroy & Muysken, 1995) which has received a great attention from researchers since the 1950s (Ibrahim, et al., 2013). Code-switching has been widely examined in the contexts of ESL (English as second language) and EFL (English as foreign language) (Weng, 2012). The main reason for investigating code-switching in classroom discourse is the fact that language classrooms are one of the social situations where speakers "share knowledge of communicative constraints and options" and, therefore, can be "said to be members of the same speech community" (Gumperz & Hymes, 1986, p. 17) and there code-switching occurs both for communication and for teaching.

Ife (2007) notifies that the use of first language (L1) in the classroom should be considered as a resource in second language acquisition (SLA), which may eventuate in bilingual competence in learners (Arnfast & Jørgensen, 2003). Using first language differs from code-switching. L1 use refers to using the students' first language to teach the target language in the classroom which differs from code-switching, i.e., the alternation between languages within a single unit such as a phrase, constituent or utterance (e.g., Poplack, 1980). In this regard, Hall and Cook (2012) explain Monolingual Teaching as a notion that "a language is best taught without reference to another language" (p. 273) and Bilingual Teaching as a notion that "use should be made of a language the student already knows" (p. 274).

As stated by Shin and Milroy (2000), "Codeswitching is used as an additional resource to achieve particular conversational goals in interactions with other bilingual speakers." (p. 351). Thinking about the topic of code-switching evokes the ideas of transferring and compensating communicative needs that provide an opportunity for learners to develop their language when they cannot express themselves because of inability in target language (Mirhasani & Jafarpour, 2009).

As attested by Merritt et al. (1992), linguistic insecurity, that is, the difficulty in relating new concepts, and socializing, i.e., indicating solidarity and intimacy, are the reasons for conducting code-switching in L2 classes. Also Flyman-Mattsson (1997), Flyman-Mattsson and Burenhult-Mattsson (1999), comment on topic switching and affective functions as intentions of using code-switching in educational context. Depending on what is reported in Nazeri et al. (2020), "helping listener with better understanding, clarification, and checking comprehension are the most important motivational determinants for code-switching" (p. 151).

In a study conducted by Mirhasani and Jafarpour (2009) on 60 low-intermediate students assigned to the control and experimental groups, using a teacher-made achievement test, it was concluded that code-switching can be used as a technique to enhance students' speaking ability. Martinez (2010) investigated the significance of code-switching between Spanish and English and stated that code-switching enhanced educational literacy and helped the students manage their conversations. Later, Mokgwathi and Webb (2013) did a research in Botswana, a country in the center of Southern Africa, and declared that code-switching increases the students' participation and comprehension. In their study with two pre-intermediate classes of an English language institute, Bozorgian and Fallahpour (2020) found out that "teachers and students resorted to the first language as an important cognitive and pedagogical tool"; moreover, "teachers maintained that using the students' first language supports second/foreign language learning and teaching processes in the pre-intermediate levels" (p. 2).

What is noteworthy to this extent is the differences between teachers' and students' code-switching in TEFL. As claimed by Nazeri (2020), students code-switch in pair or group work and in daily conversations to clarify the meaning and structure of language for themselves; moreover, to compensate their lacks and inability in target language. However, teachers mostly try to avoid code-switching since they aim to make students competent both linguistically and communicatively. As a result, the use of code-switching in student-student and student-teacher

interactions is more than teacher-student interactions (Nazeri, et al., 2020). Also Bozorgian and Fallahpour (2015) investigated the amount and purposes of L1 use in EFL classrooms and concluded that EFL teachers used a limited amount of L1 in the EFL classrooms with the purpose of improving their teaching and the students' learning. In addition, they found out that using L1 should be included in the classroom syllabi because it facilitates students' learning in EFL classrooms.

After all, the definite presence and impact of code-switching in the process of language teaching is undeniable and "the use of native language is so compelling that it emerges even when policies and assumptions mitigate against it" (Lucas & Katz, 1994, p. 558). What is important then is the kind of impact it leaves in this process. Connecting the teaching and learning processes could be regarded as the foremost impact of code-switching (Üstunel, 2016).

Nevertheless, there are two oppositions about code-switching with one side believing in its positive effects like feeling safe and expressing oneself properly (Auerbach, 1993), facilitating the process of learning, harmonizing different capacities of language competency (Brown, 2006), enriching vocabulary and grammar, relaxing learners (Ahmad & Jusoff, 2009), verifying comprehension (Moghadam, et al., 2012), making the feeling of confidence, security, motivation, friendship (Peregoy & Boyle, 2013), maintaining fluency and overcoming difficulties in lexical access (Raichlin, et al., 2018). The opposite side is concerned with classroom code-switching as a counterproductive phenomenon which plays the role of inhibition of TL learning (Eldridge, 1996); likewise, Sert (2005) asserted that code-switching causes the loss of fluency. This side regards code-switching as the sign of incorrectness (Willis, 1981) whose use could not always be effective so it is better to decrease its use as the reason (Cook, 2001) and consequently expose the students to target language in the class as much as possible (Krashen & Terrell, 1983).

It seems that the recent research tries to recognize code-switching as a helpful instrument for EFL classes. For example, Osborne (2020) mentioned the discursive and structural effects of code-switching in classrooms and Gallagher (2020) referred to the "flexible fluid and mixed views of EFL teachers in relation to the use of the L1 in the classroom" and emphasized "the need for a more explicit focus on this area in teacher development and training" (p. 1). In Johns and Steuck (2021), code-switching was identified as a unique discourse mode that is used for facilitating production and may be costly at one level but beneficial at another. After all, factors accelerating code-switching were concerned in most of the studies but whether code-switching accelerates learning has not mostly been investigated so far.

Bearing in mind the positive effects of code-switching, the EFL teachers' attitudes about this issue is worthy of attention. In his research process, Macaro (2009) explored three inter-related issues about L1 use in language classes, that is, "whether exclusive use of the target language was the best teaching approach, how this approach might affect collaborative learning, and whether exclusive use by the teacher promoted or hampered independent learning" (p. 35). In his findings, he recognized three distinct positions for teachers' theories about L1, including: 1) *Virtual Position* considering the use of second language exclusively since L2 could only be learnt through L2; 2) *Maximal Position* referring to the idea that L2 was only really learnt through the second language (L2), but it is unattainable because there exist no perfect learning conditions in language classes; 3) *Optimal Position* believing in L1 value and the role it has in learning enhancement.

By and large, based on the previous research, this study examined the use of code-switching in EFL classes from teachers' and learners' attitudes to find out what the EFL teachers' and student's attitudes are about code-switching in language classes, and whether these attitudes vary with some individual difference factors such as age, gender, social class and level of proficiency in detail.

3. METHODOLOGY

Design

According to the methodology of the previous investigations, mixed methods were mostly used in code-switching studies. Thus, the present study decided to employ triangulation mixed method design which combines both quantitative and qualitative tools in collecting, analyzing, interpreting, and reporting the data with the purpose of achieving a fuller understanding of code-switching and verifying findings.

Context of the Study

Regarding the essence of the topic of code-switching, one of the multilingual cities of East Azerbaijan Province of Iran, that is, Tabriz was selected as the context of the study where people are equipped with Turkish and Persian as their mother tongue and second language, respectively. Furthermore, English is one of the courses at high schools, and besides that, most of the students study English as a foreign language at language teaching institutes, too.

Participants

374 students and 26 teachers with Turkish as their mother tongue at English language institutes in Tabriz served as the subjects of this study. In a convenience sampling procedure, the participants were selected from five language institutes located in different regional areas of the city in terms of the social classes, that is, lower, middle, and upper-middle social class. The students, including 189 females and 185 males, from different ages including Children (5-10), Teenagers (11-20), Young Adults (21-30), Adults (31 and more), and different proficiency levels, that is, Basic, Elementary, Intermediate, High-intermediate, and Advanced, participated in this research. Also, the teachers, including 21 females and 5 males, teaching students with various levels of proficiency, participated.

Instruments and Data Collection

In the current study, the following instruments were employed for collecting data:

Questionnaire: Two separate sets of questionnaires were designed for students (Appendix A and B) and teachers (Appendix C). Both sets were based on Hymes' (1962) framework and Poplack's (1980) and Myers-Scotton's (1989) categorizations.

Observation Checklist: In order to observe the classes, an observation checklist was designed based on Motivational Orientation of Language Teaching **(**MOLT) observation scheme which was first used by Guilloteaux and Dornyei (2008), including parts based on Blom and Gumperz's (1972), Gumperz's (1982), Poplack's (1980), and Myers-Scotton's (1989) models.

For testing the feasibility, validity, and reliability of the instruments, a pilot study was conducted with 30 students and 15 teachers before starting the main phase of the research. Both students' and teachers' questionnaires were analyzed and results revealed the reliability of 0.812 and 0.629, respectively. The data related to the research questions of the present research were inferred from the related questions in the questionnaires that is, questions 20-26 in students' and questions 21-29 in teachers' questionnaires.

Data Analysis

Research Questions 1 and 3 were answered descriptively based on frequencies of replies. For analyzing the collected data to answer the Research Questions 2 and 4, the SPSS software (Version 20) was used with the purpose of applying Chi-square Test.

4. RESULTS

The main focus of the current paper was to investigate teachers' and students' attitudes regarding code-switching in Iranian EFL classes. The results of analysis related to each research question are reported below:

Research Question 1: Learners' Attitudes about Code-switching

Students' Answers

The data related to learners' attitudes about code-switching can be discussed once in general and then based on students' age, gender, and social class. The students' answers in general are provided in Table 1. The questions applying to the research reported here in students' questionnaires are 20-26, which were answered on a Likert scale of 5 components (SA= strongly agree, A= agree, SD= strongly disagree, D= disagree, and N= Neutral). Due to space limits, the frequencies and percentages for neutral responses have not been included here.

According to the Table 1, 78.7% of the students strongly agreed with practicing English through using code-switching all the time. In other words, they assume that code-switching from their L1, i.e., Turkish, or L2, i.e., Persian to English, or vice versa, leave them in a better position to learn English. About 80.5% believed that they could understand better when the teacher code-switched to Turkish or Persian, and 84.2% reported that the teachers at school usually teach all lessons in Persian and sometimes in Turkish.

Almost 59% of the students (strongly) disagreed on losing confidence and being discouraged when talking and participating in class activities because of being ridiculed or not being allowed to use code-switching. Also almost half of them (strongly) disagreed on not being allowed to code-switch to Turkish or Persian in class.

Table 1: Students' Attitudes about Code-switching

	SA+A Frequency (N)	Percentage (%)	SD+D Frequency (N)	Percentage (%)
I am ridiculed by classmates and lose confidence when I codeswitch.	94	25.1	222	59.3
Code-switching helps me practice English all the time.	294	78.7	36	9.6
The teacher usually codeswitches to Turkish or Persian during teaching.	189	50.5	126	33.7
I can understand better when the teacher codeswitches to Turkish or Persian during teaching.	301	80.5	38	10.2
We are not allowed to codeswitch to Turkish or Persian in class.	119	31.8	169	45.2
The teachers at school usually teach all lessons in Persian and sometimes in Turkish.	315	84.2	30	8.0
Since we are not allowed to codeswitch to Turkish or Persian I am not confident to talk and stay silent.	95	25.4	224	59.9

Table 2: Chi Square Test for the Effect of Students' Gender, Age, and Social Class

Characteristics		Gender			Age				Social Class		
		Girls	Boys	Children.	Teenagers	Youths	Adults	Upper-middle	Middle	Lower
20	I am ridiculed by classmates and lose confidence when I codeswitch.	Sig.	.028		.043				.001		
		SA+A%	27.5	22.7	25	26.3	13.4	26.6	26	27.6	19.3
		SD+D%	56.6	62.2	57.5	58.8	66.7	60	63.8	50.3	69.3
21	Code-switching helps me practice English all the time.	Sig.	.001		.000				.004		
		SA+A%	83.6	73.5	70	81.7	56.7	86.6	82.7	81.7	67.1
		SD+D%	3.7	15.6	27.5	7.3	10	6.7	11.8	6.9	11.3
22	The teacher usually codeswitches to Turkish or Persian during teaching.	Sig.	.437		.014				.007		
		SA+A%	48.6	52.4	75	48.1	46.7	40	59.1	42.8	52.3
		SD+D%	32.8	34.6	22.5	34.6	33.3	46.7	29.1	38.3	31.8
23	I can understand better when the teacher codeswitches to Turkish or Persian during teaching.	Sig.	.133		.000				.052		
		SA+A%	78.3	82.7	80	81	66.7	100	82.7	83.6	71.6
		SD+D%	13.2	7	20	20	7.9	23.4	11	6.3	16
24	We are not allowed to codeswitch to Turkish or Persian in class.	Sig.	.027		.000				.001		
		SA+A%	36	27.6	50	30.8	20	26.7	40.9	30.2	21.6
		SD+D%	38.1	52.4	42.5	44.3	63.4	33.4	40.2	43.4	55.6
25	The teachers at school usually teach all lessons in Persian and sometimes in Turkish.	Sig.	.160		.007				.010		
		SA+A%	88.3	80	87.5	85.1	80	44.7	87.4	81.1	85.6
		SD+D%	4.7	11.3	7.5	6.9	16.7	13.3	5.5	10.7	6.8
26	Since we are not allowed to codeswitch to Turkish or Persian, I am not confident to talk and stay silent.	Sig.	.515		.282				.024		
		SA+A%	25.9	24.9	37.5	25.2	6.6	33.3	28.3	25.8	20.5
		SD+D%	57.7	62.2	50	60.4	73.3	46.7	58.2	55.4	70.5

Research Question 2: Learners' Gender, Age and Social Class

The results related to possible differences in learners' attitudes about code-switching in their EFL classes which can be attributed to learners' age, gender, and social class are presented in Table 2.

Gender

The amount of Asymp. Sig (2-tailed) is less than .05 just in two cases. Although less than half of the participants, both boys and girls, agreed or strongly agreed on being ridiculed by classmates and losing confidence when they code-switched, the number of girls agreeing on this case is a bit more than boys. Also, girls (strongly) agreed about not being allowed to code-switch to Turkish or Persian in class.

Age

No differences were observed between age groups in terms of Question 26 that concerned not being allowed to code-switch in Turkish or Persian and not being confident to talk and stay silent. However, in all other cases significant differences were observed between age groups. Less than half of the students-- almost 26% of the adults and teenagers and 25% of the children (strongly)

agreed that they were ridiculed by classmates and lost their confidence when they codeswitched. In this case, the least percentage is related to the young adults. 86.6% of the adults, 81.7% of the teenagers, 70% of the children, and 56.7% of the young adults reported that code-switching helps them practice English all the time. 75% of the children and less than half of the other age groups, i.e., 48.1% of the teenagers, 46.7% of the youths, and 40% of the adults, agreed or strongly agreed that the teachers usually codeswitch to Turkish or Persian during teaching. All of the adults -- 81% of the teenagers, 80% of the children, and 66.7% of the youths stated that they can understand better when the teacher codeswitches to Turkish or Persian during teaching. Half of the children, 30.8% of the teenagers, 26.7% of the adults, and 20% of the youths reported they are not allowed to codeswitch in Turkish or Persian in class. More than 80% of the children, the teenagers, and the youths reported that the teachers at school or university usually teach all lessons in Persian and sometimes in Turkish.

Social Class

According to the amount of Asymp. Sig (2-tailed) which should be less than .05, there are significant differences between three groups of social class with regard to all questions except Question 23. Less than half of the participants in all social classes agreed or strongly agreed on being ridiculed by classmates and lose confidence when they code-switched; however, among them the lower class has reported the least and the upper-middle and the middle class are almost identical. Nearly 82% of the upper-middle and the middle class and 67.1% of the lower class believed that code-switching helped them practice English all the time. 59.1% of the upper-middle class and more than half of the lower class reported that the teachers usually codeswitched to Turkish or Persian during teaching but this frequency is 42.8% in the middle class. The frequency of not being allowed to codeswitch to Turkish or Persian in class is from more to less in the upper-middle, middle, and lower classes, respectively. The lower class has the lowest frequency in being allowed to codeswitch to L1 or L2. In all social classes with order of the upper-middle, lower, and middle class, the teachers at schools usually teach all lessons in Persian and sometimes in Turkish. About not being allowed to codeswitch to Turkish or Persian and not being confident to talk and staying silent, less than half and almost 20% to 28% agreed or strongly agreed. The highest frequency in this case is related to upper-middle and the lowest belongs to the lower class.

Research Question 3: Teachers' Attitudes about Code-switching

Teachers' Answers

The teachers' answers regarding their attitudes about code-switching in their classes in general are provided in Table 3. Furthermore, teachers' answers classified based on the level they teach are provided in Table 4. The questions applying to the research reported here in teachers' questionnaires are 21-29, which were answered on a Likert scale of 5 components (SA= strongly agree, A= agree, SD= strongly disagree, D= disagree, and N= Neutral). Due to space limits, the frequencies and percentages for neutral responses have not been included here.

Table 3: Teachers' Attitudes about the Effect of Code-switching4

		SA+A Frequency (N) Percentage (%)		SD+D Frequency (N) Percentage (%)	
21	Code-switching eases up teaching method.	11	42.3	15	57.7
22	Code-switching wastes time in the classroom.	13	50	9	34.6
23	The students give positive feedback (participation, results, etc.) when I codeswitch.	9	34.6	13	50.0
24	The students still get confused when I codeswitch.	16	61.5	4	15.3
25	Code-switching does not promote English speaking environment.	10	38.4	10	38.5
26	I'm being asked to codeswitch by my students.	10	38.4	14	53.9
27	The students become fully dependent on code-switching for better understanding.	8	30.8	15	57.7
28	Using code-switching leads to the weakness of the students' English.	8	30.7	18	69.3
29	Using code-switching strengthens the students' English learning.	14	53.9	8	30.8

According to the Table 3, 42.3% of the teachers (strongly) agreed that code-switching eases up teaching method but 57.7% of them (strongly) disagreed about it. Almost half of the teachers stated that code-switching wastes time of the classroom. Also 50% reported they do not get positive feedback (participation, results, etc.) from the students when they codeswitch. In other words, they declared that when they codeswitch, the students' participation in activities, and consequently, their grades decrease. 61.5% of teachers(strongly) agreed that the students still get confused when they codeswitch. 53.9% of them declared they are not being asked to codeswitch by the students. About 57.7% of the subjects stated that the students do not become fully dependent on code-switching for better understanding, and almost 70% believed that using code-switching does not lead to the weakness of the students' English. Moreover, 53.9% strongly agreed or agreed that using code-switching strengthens the students' English.

Research Question 4: The Effect of Students' Level of Proficiency on Teachers' Attitudes

Among all teacher-related variables potentially moderating teachers' attitudes about code-switching, the Level of the Class they were teaching was hypothetically predicted to be relevant. The results of the analysis regarding the differences between teachers teaching at each of the 5 levels of language classes in terms of their attitudes about code-switching have been presented in Table 4.

The results of Chi Square test provided in Table 4 indicated no differences between the attitudes of teachers resulting from the level at which they were teaching since the amounts of Asymp. Sig (2-tailed) was above .05 in all cases.

Table 4: Chi Square Test for the Impact of Teachers' Teaching Level

	Characteristics		Basic	Elementary	Intermediate	High-Intermediate	Advanced
21	Code-switching eases up teaching method.	Sig.	.761				
		SA+A%	57.1	100	55.5	0	60
		SD+D%	28.6	0	22.2	50	20
22	Code-switching wastes time in the classroom.	Sig.	.650				
		SA+A%	42.9	0	44.4	50	20
		SD+D%	28.6	100	55.5	0	60
23	The students give positive feedback (participation, results, etc.) when I codeswitch.	Sig.	66.7				
		SA+A%	100	66.7	22.2	0	40
		SD+D%	0	0	66.7	50	40
24	The students still get confused when I codeswitch.	Sig.	.562				
		SA+A%	0	0	22.2	50	20
		SD+D%	71.5	100	55.5	0	60
25	Code-switching does not promote English speaking environment.	Sig.	.419				
		SA+A%	28.6	33.3	44.4	100	20
		SD+D%	28.6	66.7	33.3	0	60
26	I'm being asked to codeswitch by my students.	Sig.	.077				
		SA+A%	42.9	100	77.8	50	0
		SD+D%	57.2	0	22.2	50	60
27	The students become fully dependent on code-switching for better understanding.	Sig.	.631				
		SA+A%	57.2	100	55.6	50	40
		SD+D%	28.6	0	22.2	50	60
28	Using code-switching leads to the weakness of the students' English.	Sig.	.296				
		SA+A%	57.2	33.3	100	100	40
		SD+D%	14.3	33.3	0	0	20
29	Using code-switching strengthens the students' English learning.	Sig.	.475				
		SA+A%	57.2	100	44.4	50	60
		SD+D%	14.3	0	33.3	0	20

5. RESULTS OF OBSERVATION

For analyzing the observation checklists, all parts were coded and imported to SPSS (Version 20). Distribution and frequency of data, along with correlational analysis were computed for checklists of 22 classes. We analyzed data through coding, identifying themes, interpreting, and providing meaning through inferential analysis. The results indicated that the teachers had two approaches about code-switching. Most of them tried to avoid code-switching which usually resulted in wasting class time for providing abstract and ambiguous explanations. However, the teachers who did not have prohibition for code-switching could accelerate their teaching and students' learning by saving time and directing students' attention to what was the main point. These classes were of two kinds: a) The teacher could manage the class after code-switching and

continued their teaching in English; b) classes where code-switching was not used in restricted way; hence, shifted into L1 or L2 use. This shift could create some hidden problems in students' learning in longtime.

6. DISCUSSION

The main concern in this research was to investigate how Iranian EFL teachers and learners perceived code-switching in FL learning, and whether these attitudes varied according to some individual difference and contextual factors. Clarifications on this can contribute to outlining a detailed 'effectiveness scheme' for code-switching in EFL classes. The findings attest to the relevance of affective, social and contextual variables to the way teachers and students encounter code-switching. Therefore, most of the students asserted that they could understand better when the teachers codeswitched to Turkish or Persian during teaching. This refers to the functions called *interjections* (Gumperz, 1982) and *reiteration* (Eldridge, 1996) both of which mean to use code-switching to clarify and conform the message for better understanding.

Students expressed that they could practice English when they used code-switching and half of them reported that the teachers codeswitched during teaching. The reasons for which the teachers codeswitched could be: 1) to provide L1 equivalents in target language (equivalence); 2) to fill the conversational gaps (floor-holding); and 3) to manage the clash use of language (conflict) (as in Eldridge, 1996); 4) to evaluate the comprehension; 5) to affirm and stimulate the participation; and 6) to manage the classroom (Myers-Scotton, 1993).

Mostly the students mentioned that using code-switching did not make them lose their confidence. This idea refers to what has been proposed by Peregoy and Boyle (2013) who declared that code-switching facilitates teaching and learning by making the feeling of confidence, security, motivation, and friendship. Also Gomez (2014) stated that using the students' native language makes them feel their L1 identities are valued and improves their learning.

In all social classes especially with the upper-middle social class with the highest frequency, a number of students, particularly girls declared that they were not allowed to codeswitch to Turkish or Persian in class. Also about preventing the use of code-switching, the results revealed that girls were prevented from code-switching more than boys. According to Üstunel (2016), sometimes the students tend to switch language because their language proficiency is not the same as their peers or is not equal to the teachers' mastery. Using code-switching prevents miscommunication and as Moore (2010) argued, the students may codeswitch to clarify what is being talked. Therefore, preventing students to use code-switching can affect their language learning negatively which may result in lower language proficiency.

In multilingual pedagogy, the individuals deal with different languages, i.e., their mother tongue, second language, and maybe other languages. In the context of Iran, Persian is the language of education. Meanwhile, in Tabriz, Northwest of Iran, students deal with Turkish as L1 and Persian as the official language, and the majority of them declared that all courses at schools were usually taught and learned in Persian as well. In addition to education, the language of media is Persian too. So the multilingual Turkish speakers are always dealing with Persian in reading books, writing letters, and watching TV. Beside Persian, individuals learn Arabic language at school, which is their religious language. They learn Arabic vocabulary and structures and unconsciously use some Arabic expressions in their daily conversations.

Subsequently, when these multilingual individuals, who already have Turkish, Persian, and Arabic structures in mind, attend English classes, and start to deal with a new language and culture, they try to put what they learn in the form of the languages they are already equipped with (Nazeri, 2020). Hymes (1962) has focused on communicative functions of code-switching and suggested that one of the functions of classroom code-switching is *poetic functions* which means to insert

some jokes, stories, and poetic quotations in order to add a sense of humor. Based on the results of the observation checklists used in this study, they mostly translate what they hear into Turkish or Persian, they use lots of Arabic expression like Ya *Allah, Masha Allah, Insha Allah*, etc. or Turkish expressions like *Vay Dada!* for joking and making fun or showing their wonder and surprise, in their conversations. Most of the students especially the adults, the teenagers, and the children in all social classes believe that code-switching helps them practice English all the time. Moreover, in all social classes especially the upper-middle and the lower social class, mostly children have reported that the teachers usually codeswitch to L1 or L2 during teaching.

Although the majority of the teachers believe that the students are still confused when they codeswitch, the teenagers, children, and young adults have mostly reported that they can understand better when the teacher codeswitches. The same finding was reported by Al-Qaysi (2016) that educators codeswitched in their lectures to help the students understand better. Besides, the teachers in the current study believed that code-switching strengthens the students' English learning. This is in line with Ahmad and Jusoff (2009), who found that teachers' code-switching was an effective teaching strategy when dealing with low English proficient learners. Therefore, using code-switching enhances effective learning foreign language (Akynova, et al., 2012).

Although some of the teachers argued that code-switching eases up teaching, as mentioned in Uys and Van Dulm, (2011), they asserted that it wastes time, does not promote English speaking environment, and when using code-switching they do not get positive feedback from the students, and thus their common belief is that *the best English teachers just teach in English*. However, the teachers say that they are being asked to codeswitch which they think may lead to the weaknesses of the students since they may become fully dependent on code-switching.

One important issue that should be considered all the time is the fact that the teachers should distinguish between code-switching and using first or second language. The students cannot avoid Turkish identity, Persian thinking, and Arabic expressions which have been mixed with their language in a way that they use them unconsciously. Thus, using pure L1 or L2 in teaching target language can be problematic and can impede learning (Nazeri, 2020). On the contrary, code-switching can accelerate learning. Students and teachers mostly reported favorable attitudes towards code-switching, and also the observation results indicated that code-switching can save much of class time. This is in accordance with the dominant literature on the effectiveness of code-switching. Üstunel (2016) believes code-switching connects the teaching and learning process. Furthermore, Enama (2016) states that the target language learning should take place together with first language. There should not be any burden of employing L1 because it serves the precise function to the students in the class (Enama, 2016).

7. CONCLUSION

Alongside providing a brief history about code-switching and its' motivational determinants and attitudes toward its use in language classrooms, the present study attempted to indicate the relative differences between teachers' and students' code-switching and its role in acceleration of learning a foreign language. In spite of the fact that the results disclosed almost the cynical perspectives of teachers about code-switching i.e., they think that code-switching may impede learning, both students' questionnaire and the observation reports demonstrated that code-switching accelerates not only learning but also teaching English.

Since the students reported that teachers at school usually teach all lessons in Persian and sometimes in Turkish, they deal with L1 and L2 most of the time so the role and impact of these languages in FL learning is incontestable. Likewise, students declare that they practice English through code-switching and can understand better when the teacher codeswitches.

Using code-switching saves the time of the class and prevents wasting the time for explaining the subject matters with abstract definitions which are completely incomprehensible to students. In other words, by introducing code-switching to our teaching methodology, we as the teachers provide ourselves with a teaching strategy which benefits our class as well as our students. It is irrefutable that everyone's identity and culture is attached to them and we, whether being averse or not, cannot abnegate this priority. Thus, instead of denying code-switching, rebuking ourselves for using it consciously or unconsciously during teaching, and reproaching our students for code-switching to their mother tongue, it is the time that we should recognize code-switching as an aid that assists on the teaching-learning process.

The current study attempted to investigate the role of code-switching in acceleration of foreign language learning in a multilingual context. It could be replicated in any context where any language is taught as second or foreign language and could look into more cases by increasing the number of participants, conducting individual interviews with the teachers and the students, increasing the number of classes to observe, doing an ethnographic research, doing the same research in schools where a foreign language is taught, and finally investigating the effect of code-switching in TEFL or English language proficiency of the students by conducting research with experimental designs.

Like most of the other studies based on self-reporting, the current study suffered from the data reliability limitation. Although we tried to minimize this deficiency by increasing the number of participants, it did not become possible in the case of teachers. A more realistic view of students and teachers' attitudes regarding instances of effective code-switching requires a more in-depth and closer probe into their inclinations during the teaching-learning process. Complimented with the more empirical research designed to compare the teaching schemes with and without integration of code-switching can pave the path for a better understanding of the effectiveness of code-switching in L2 classes.

References

Ahmad, B., & Jusoff, K. (2009). The teachers' code-switching in the classroom instructions for low English proficient learners. *English Language Teaching, 2*(2), 49-55. https://doi.org/10.5539/elt.v2n2p49

Al-Qaysi, N. J. M. (2016). Examining the students' and educators' attitudes towards the use of code-switching within higher educational environments in Oman. (Unpublished master's thesis). The British University in Dubai, United Arab Emirates.

Akynova, D., Zharkynbekova, S., & Aimoldina, A. (2012). Code-switching and second language acquisition in the educational space of Kazakhstan. *GSTF International Journal of Law and Social Sciences, 2*(1), 224-230. https://doi.org/10.5176/2251-2853_2.1.84

Arnfast, J. S., & Jørgensen, J. N. (2003). Code-switching as a communication, learning, and social negotiation strategy in first-year learners of Danish. *International Journal of Applied Linguistics, 13*(1), 23-53. https://doi.org/10.1111/1473-4192.00036.

Auerbach, E. (1993). Reexamining English only in the ESL classroom. *TESOL Quarterly, 27*(1), 9-32. https://doi.org/10.2307/3586949.

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language, 95*, 173-198. https://doi.org/10.1016/j.jml.2017.04.002.

Bozorgian, H., & Fallahpour, S. (2020). Exploring L1 use in English as foreign language classrooms through activity theory. *Journal of Teaching Language Skills (JTLS), 39*(3.1), 1-35. https://doi.org/10.22099/jtls.2021.39025.2912

Bozorgian, H., & Fallahpour, S. (2015). Teachers' and students' amount and purpose of L1 use: English as foreign language (EFL) classrooms in Iran. *Iranian Journal of Language Teaching Research, 3*(2), 67-81. https://doi.org/10.30466/ijltr.2015.20390

Blom, J., & Gumperz, J. J. (1972). Social meaning and linguistic structures: Code-switching in Northern Norway. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 407-434). New York: Holt, Rinehart and Wilson.

Brown, K. (2006). *Encyclopaedia of language & linguistics*. Elsevier.

Cook, V. (2001). Using the first language in the classroom. *The Canadian Modern Language Review, 57*(3), 402-415. https://doi.org/10.3138/cmlr.57.3.402

Eldridge, J. (1996). Code-switching in a Turkish secondary school. *ELT Journal, 50*(4), 303-311. https://doi.org/10.1093/elt/50.4.303.

Enama, P. R. B. (2016). The impact of English-only and bilingual approaches to EFL instruction on low-achieving bilinguals in Cameroon: An empirical study. *Journal of Language Teaching and Research, 7*(1), 19-30. https://doi.org/10.17507/jltr.0701.03.

Flyman-Mattsson, A. (1997). Communication strategies in French as a foreign language. *Working Papers, 46*, 57-75. Department of Linguistics, Lund.

Flyman-Mattsson, A., & Burenhult-Mattsson, N. (1999). Code-switching in second language teaching of French. *Working Papers, 47*, 59-72. Department of Linguistics, Lund University.

Gallagher, F. (2020). Considered in context: EFL teachers' views on the classroom as a bilingual space and codeswitching in shared-L1 and multilingual contexts. *System, 91*, Article 102262. https://doi.org/10.1016/j.system.2020.102262

Gomez, A. (2014). How do native Spanish speakers engage with and respond to bilingual instructional strategies in the ESL classroom? Retrieved from http://bilingualstrategiesinesl.weebly.com/

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Harvard University Press.

Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly, 42*(1), 55-77. https://doi.org/10.1002/j.1545-7249.2008.tb00207.x

Gumperz, J. (1982). *Discourse strategies*. Cambridge University Press.

Gumperz, J., & Hymes, D. H. (1986). *Directions in sociolinguistics: The ethnography of communication*. Basil Blackwell.

Hymes, D. (1962). The ethnography of speaking. In T. Gladwin (Ed.).

Kim, M. (2008). Are cases of unconventional code-switching useless? *The Panels, 44*(2), 139-149.

Ibrahim, E., Shah, M., & Armia, N. (2013). Code-switching in English as a foreign language classroom: The teachers' attitudes. *English Language Teaching, 6*(7), 139-150. https://doi.org/10.5539/elt.v6n7p139

Ife, A. (2007). A role for English as lingua franca in the foreign language classroom. In A. Soler & S. Jordà (Eds.), *Intercultural language use and language learning* (pp. 79-100). Springer.

Johns, M. A., & Steuck, J. (2021). Is code-switching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech. *System*, 211, Article 104634. https://doi.org/10.1016/j.cognition.2021.104634

Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. The Alemany Press.

Lin, A. M. Y. (2013). Classroom code-switching: Three decades of research. *Applied Linguistics Review, 4*(1), 195-218. https://doi.org/10.1515/applirev-2013-0009

Lucas, T., & Katz, A. (1994). Reframing the debate: The roles of native language in English-only programs for language minority students. *TESOL Quarterly, 28*(3), 537-561. https://doi.org/10.2307/3587307

Macaro, E. (2009). Teacher use of codeswitching in the second language classroom: Exploring 'optimal' use. In M. Turnbull & J. Dailey-O'Cain (Eds.), *First language use in second and foreign language learning* (pp. 35-49). Multilingual Matters.

Martinez, R. (2010). Spanglish as literacy tool: Toward an understanding of the potential role of Spanish-English code-switching in the development of academic literacy. *Research in the Teaching of English, 45*(2), 124-149. Retrieved from http://www.jstor.org/stable/40997087

Merritt, M., Cleghorn, A., Abagi, J. O., & Bunyi, G. (1992). Socializing multilingualism: Determinants of code-switching in Kenyan primary classrooms. In C. M. Eastman (Ed.), *Code-switching* (pp. 103-121). Multilingual Matters Ltd.

Miles, R. (2004). Evaluating the use of L1 in the English language classroom. (Unpublished master's dissertation). University of Birmingham, UK.

Milroy, L., & Muysken, P. (1995). Introduction: Code-switching and bilingualism research. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching* (pp. 1-14). Cambridge University Press.

Mirhasani, A., & Jafarpour Mamaghani, H. (2009). Code-switching and Iranian EFL learners' oral proficiency. *Journal of Teaching English as a Foreign Language and Literature, 1*(2), 21-31. Retrieved from https://www.sid.ir/en/journal/ViewPaper.aspx?id=201515

Moghadam, S., Samad, A., & Shahraki, E. (2012). Code-switching as a medium of instruction in an EFL classroom. *Theory and Practice in Language Studies, 2*(11), 2219-2225. https://doi.org/10.4304/tpls.2.11.2219

Mokgwathi, T., & Webb, V. (2013). The educational effects of code-switching in the classroom—Benefits and setbacks: A case of selected senior secondary schools in Botswana. *Language Matters: Studies in the Languages of Africa, 44*(3), 108-125. https://doi.org/10.1080/10228195.2013.839734

Moore, D. (2010). Code-switching and learning in the classroom. *International Journal of Bilingual Education and Bilingualism, 5*(5), 279-293. https://doi.org/10.1080/13670050208667762

Myers-Scotton, C. (1993). *Social motivations for codeswitching: Evidence from Africa*. Clarendon Press.

Myers-Scotton, C. (1989). Codeswitching with English: Types of switching, types of communities. *World Englishes, 8*(3), 333-346. https://doi.org/10.1111/j.1467-971x.1989.tb00673.x

Nazeri, S., Amini, D., & Salahshoor, F. (2020). Motivational determinants of code-switching in Iranian EFL classrooms. *The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances, 8*(1), 151-173. https://doi.org/10.22049/jalda.2020.26812.1171.

Nazeri, S. (2020). The sociolinguistic study of code-switching in TEFL in Iranian English language institutes. (Master's thesis). Azerbaijan Shahid Madani University, Tabriz, Iran.

Osborne, D. (2020). Codeswitching practices from "other tongues" to the "mother tongue" in the provincial Philippine classroom. *Linguistics and Education, 55*, Article 100780. https://doi.org/10.1016/j.linged.2019.100780

Peregoy, S. F., & Boyle, O. F. (2013). *Reading, writing, and learning in ESL: A resource book for teaching K-12 English learners* (6th ed.). Pearson.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics, 18*, 581-618. https://doi.org/10.1515/ling.1980.18.7-8.581.

Poplack, S. (2000). Sometimes I'll start a sentence in English y termino en español: Toward a typology of code-switching. In L. Wei (Ed.), *The bilingualism reader* (pp. 221-256). Routledge.

Raichlin, R., Walters, J., & Altman, C. (2019). Some wheres and whys in bilingual codeswitching: Directionality, motivation, and locus of codeswitching in Russian-Hebrew bilingual children. *International Journal of Bilingualism, 23*(2), 629-650. https://doi.org/10.1177/1367006918763135

Samar, R. G., & Moradkhani, S. (2014). Codeswitching in the language classroom: A study of four EFL teachers' cognition. *RELC Journal, 45*(2), 151-164. https://doi.org/10.1177/0033688214534796

Sert, O. (2005). The functions of code-switching in ELT classrooms. *The Internet TESL Journal, XI*(8). Retrieved from http://iteslj.org/Articles/Sert-CodeSwitching.html

Shin, S. J., & Milroy, L. (2000). Conversational code-switching among Korean-English bilingual children. *International Journal of Bilingualism, 4*(3), 351-383. https://doi.org/10.1177/13670069000040030401

Turnbull, M. (2001). There is a role for the L1 in second and foreign language teaching, but… *The Canadian Modern Language Review, 57*(4), 531-540. https://doi.org/10.3138/cmlr.57.4.531

Üstunel, E. (2016). *EFL classroom code-switching*. Palgrave Macmillan.

Uys, D., & van Dulm, O. (2011). The functions of classroom code-switching in the Siyanda District of the Northern Cape. *Southern African Linguistics and Applied Language Studies, 29*(1), 67-76. https://doi.org/10.2989/16073614.2011.583159

Wei, L. (2008). Research perspectives on bilingualism and multilingualism. In L. Wei & A. G. Moyer (Eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism* (pp. 3-17).

Weng, P. (2012). Code-switching as a strategy in an EFL classroom in Taiwan. *US-China Foreign Language, 10*(10), 1669-1675. https://doi.org/10.17265/1539-8080/2012.10.007

Willis, J. (1981). *Teaching English through English*. Longman.

Appendix A: Students' Questionnaire

سلام

زبان آموزان عزیز، در ابتدا از اینکه وقت ارزشمندتان را برای پاسخ دادن به این پرسشنامه اختصاص داده اید سپاسگزاریم. موضوع تحقیق درباره تغییر زبان "Codeswitching" است. هدف دانستن نظرات شما نسبت به این موضوع، بررسی دلایل استفاده از codeswitching، پررنگ تر کردن مزایا و یافتن راه حل مناسب برای معایب و مشکلات احتمالی است. این پرسشنامه برای افرادی تنظیم شده که زبان مادری شان ترکی است. از اینکه صادقانه به سوالات پاسخ می دهید متشکریم.

*CS مخفف واژه code-switching است بمعنی تغییر بین دو یا چند زبان هنگام صحبت کردن (مثلا از انگلیسی به فارسی یا ترکی).

جنسیت: مونث ☐ مذکر ☐ **سن**:........ **سطح** (level): Basic☐ Elementary☐ Intermediate☐ High☐ Advanced☐

کاملا مخالفم	مخالفم	نظری ندارم	موافقم	کاملا موافقم	سوالات	
					من در مکالماتم از CS استفاده می‌کنم.	1
					من معمولا از کلمات انگلیسی Yes, No, Thank you درمکالماتم استفاده می‌کنم.	2
					در کلاس زبان هنگام انجام کار گروهی با هم گروهی‌هایم از CS استفاده می‌کنم.	3
					من در کلاس زبان در آموزشگاه به زبان انگلیسی ولی خارج از کلاسبه زبان ترکی یا فارسی با مدرس صحبت می‌کنم.	4
					CS مهارت‌های ارتباطی من را تقویت می‌کند.	5
					CS به رشد مهارت‌های زبانی من کمک می‌کند.	6
					استفاده از CS نشان می‌دهد که من فرد باسوادی هستم و دانش زیادی دارم.	7
					استفاده از CS نشان می‌دهد که من فرد با اعتباری هستم.	8
					CS مدرس و دانش آموزان بین زبان‌های ترکی و انگلیسی و فارسی تاثیر مثبتی در یادگیری زبان دارد.	9
					موقع صحبت کردن به زبان ترکی یا فارسی گاها از بعضی کلمات انگلیسی استفاده می‌کنم چون آن کلمات معادل ترکی یا فارسی ندارند و یا من آن لحظه بخاطر نمی‌آورم.	10
					موقع صحبت کردن به زبان انگلیسی گاها نمی‌توانم بعضی کلمات را بخاطر بیارم و یا بلد نیستم به همین خاطر از کلمات ترکی یا فارسی استفاده می‌کنم.	11
					CS از ترکی یا فارسی به انگلیسی یا برعکس به من کمک می‌کند تا کلمات جدید را راحت‌تر انتقال دهم.	12
					بخاطر پیچیدگی بعضی از لغات در زبان مادریام (ترکی) یا زبان دومم (فارسی) از CS استفاده می‌کنم یعنی از ترکی یا فارسی مدام به انگلیسی CS می‌کنم.	13
					CS از فارسی و ترکی به انگلیسی و برعکس به من کمک می‌کند نظرات و احساساتم را راحت‌تر بیان کنم.	14
					CS به من کمک می‌کند تا بهتر متوجه شوم.	15
					من اغلب فقط یک کلمه در جمله انگلیسی را codeswitch می‌کنم.	16
					من اغلب یک جمله کامل را codeswitch می‌کنم.	17
					من اغلب یک عبارت را codeswitch می‌کنم.	18
					من اغلب در داخل کلمه codeswitch می‌کنم. (مثال: apple ها).	19

					وقتی در کلاس زبان CS انجام می‌دهم مورد تمسخر همکلاسی‌هایم قرار می‌گیرم.	**20**
					استفاده از CS از ترکی یا فارسی به انگلیسی به من کمک می‌کند تا زبان انگیسی را در طول روز تمرین کنم.	**21**
					مدرس در کلاس از CS از انگلیسی به فارسی یا ترکی استفاده می‌کند.	**22**
					وقتی مدرس از CS از انگلیسی به فارسی یا ترکی استفاده میکند من درس را بهتر متوجه می‌شوم.	**23**
					ما در کلاس زبان در آموزشگاه اجازه نداریم از CS از انگلیسی به فارسی یا ترکی استفاده کنیم.	**24**
					در مدرسه یا دانشگاه معلم ها یا اساتید درس را بیشتر به زبان فارسی و گاها به ترکی توضیح می‌دهند.	**25**
					من بخاطر اینکه اجازه استفاده از CS در کلاس نداریم و باید انگلیسی صحبت کنیم اعتماد بنفس ندارم که صحبت کنم و معمولا ساکتم.	**26**

Appendix B: Little Students' Questionnaire

به منظور تسهیل در امر پاسخگویی زبان آموزان کم سن و سال، محقق سوالات پرسشنامه را برای زبان آموزان می خواند و آنها شکلک مورد نظرشان را علامت می‌زنند.

جنسیت: مونث☐   مذکر☐   **سن**: ....

سطح Advanced☐   High☐   Intermediate☐   Elementary☐   Basic☐ (level)

کاملا مخالفم	مخالفم	نظری ندارم	موافقم	کاملا موافقم	
😩	🙁	😐	🙂	😍	1
😩	🙁	😐	🙂	😍	2
😩	🙁	😐	🙂	😍	3
😩	🙁	😐	🙂	😍	4
😩	🙁	😐	🙂	😍	5
😩	🙁	😐	🙂	😍	6
😩	🙁	😐	🙂	😍	7
😩	🙁	😐	🙂	😍	8
😩	🙁	😐	🙂	😍	9
😩	🙁	😐	🙂	😍	10
😩	🙁	😐	🙂	😍	11
😩	🙁	😐	🙂	😍	12
😩	🙁	😐	🙂	😍	13
😩	🙁	😐	🙂	😍	14
😩	🙁	😐	🙂	😍	15
😩	🙁	😐	🙂	😍	16
😩	🙁	😐	🙂	😍	17
😩	🙁	😐	🙂	😍	18
😩	🙁	😐	🙂	😍	19
😩	🙁	😐	🙂	😍	20
😩	🙁	😐	🙂	😍	21
😩	🙁	😐	🙂	😍	22
😩	🙁	😐	🙂	😍	23
😩	🙁	😐	🙂	😍	24
😩	🙁	😐	🙂	😍	25
😩	🙁	😐	🙂	😍	26

Appendix C. Teachers' Questionnaire

Dear Lecturer/Instructor,

Thank you for participating in this questionnaire. You are going to deal with questions about Code-switching (refers to alternating between one or more languages). This questionnaire is designed to find out types of code-switching, motivational factors of code-switching and your opinion about code-switching. Please answer the questions honestly.

Gender: Female☐ Male☐   **Edu. Level**:   MS☐   BA☐   MA☐ PhD☐
Teaching experience: …………years
Level you teach: Basic☐   Intermediate☐   High☐   Advanced☐

	Questions	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
1	I use English for teaching.					
2	I use Persian for teaching.					
3	I use Turkish for teaching.					
4	I usually code switch from English to Turkish/Persian when I teach.					
5	I use code-switching to give tasks.					
6	I use code-switching to translate and clarify difficult vocabulary.					
7	I use code-switching to boost students to participate in class activities.					
8	I often codeswitch to English when I am talking in Turkish/ Persian.					
9	I only use English when I feel I'm being observed. Mostly, I teach in Persian or Turkish.					
10	I only codeswitch to Turkish or Persian when teaching new terms.					
11	I only codeswitch to Turkish or Persian when my students are confused.					
12	I teach better when I codeswitch.					
13	Code-switching saves time in teaching.					
14	Codeswitch simplifies teaching.					
15	Students understand better when I codeswitch.					
16	I feel more comfortable when I communicate with my students in language other than English.					
17	I often use code-switch one word within an English sentence.					
18	I often codeswitch the complete sentence to Turkish or Persian.					
19	I use code-switching for tag phrases.					
20	I often use code-switching within the word.					
21	Code-switching eases up teaching method.					
22	Code-switching wastes time in the classroom.					
23	Students give positive feedback when I codeswitch.					
24	Students still get confused when I codeswitch.					
25	Code-switching does not promote English speaking environment.					
26	I'm being asked to codeswitch by my students.					
27	Students become fully dependent on code-switching for better understanding.					
28	Using code-switching leads to the weakness of students' English.					
29	Using code-switching strengthens students' English.					
30	Code-switching is important in teaching any subject.					
31	Code-switching is necessary in Iranian context.					
32	Code-switching can be planned in teaching.					
33	Code switching should be avoided.					

34	I think teaching courses only in English language is beneficial for students.					
35	Teaching courses in English and a language other than English makes it easy for students to understand.					

Questioning and Responding in Target Language Use Situation Tasks: A Think-aloud Study

Mina Madadi[1] & Reza Rezvani[1,*] iD

[1] Yasouj University, Yasouj, Iran[†]

Abstract

The present study reports a qualitative study aiming to explore whether cognitive processes underlying responding and questioning in English are cognitively analogous. It further sought to discover the latent differences between responding and questioning cognitive processes in Target Language Use Situation tasks. To this end, 20 Iranian general IELTS applicants from two different institutes in Shiraz, with two different language proficiency levels (intermediate and advanced) participated in the study. They were administered a normal oral IELTS responding and a reverse questioning task. Articulated Thoughts in a Simulated Situation (ATSS) paradigm as a think-aloud approach was used to collect qualitative data during task completion. The applicants' recorded voices during task completion were transcribed and analyzed to examine the potential differences between responding and questioning cognitive processes. The analysis of the qualitative data through ATSS paradigm, in general, indicated that the cognitive processes underlying these two processes are not exactly parallel. The applicants tended to be more accurate and fluent during responding tasks. More specifically, they had fewer and shorter pauses and generated more intelligible and comprehensible productions by committing fewer grammatical errors in both proficiency levels. The paper discusses the findings and the implications for second language learners, teachers, and test developers.

1. INTRODUCTION

Assessment plays a key role in any educational system. It is defined as "Any systematic method of obtaining information from tests and other sources, used to draw inferences about characteristics of people, objects, or programs" (AERA, 1999, p. 172). Essentially, a language test is a procedure for eliciting implicit knowledge through what we can observe, and from which we can infer the amount and type of language knowledge which we cannot directly observe (Douglas, 2010). Assessments developed and used locally are likely to hold lower stakes than large-scale tests which

are "conducted on a regional, national, or international scale involving large student populations" (Simon, Ercikan, & Rousseau, 2012, p.1)

The key to large-scale assessment is uniformity in development and administration (Kunnan, 2009; Wendler & Walker, 2006). This uniformity and systematicity is both an asset (Kunnan, 2009) and may lead to imprecision and potential misinterpretation (Fox, 2008; Read, 2009) when used across time, regions, administrations, and examinees. Developers of large-scale tests are, thus, accountable for their products, and test-users are held responsible to ensure that appropriate interpretation and use of test scores are made (Chalhoub-Deville & Turner, 2000). The higher the stakes a test takes, the more considerable the demand for its validation.

Large-scale language tests have also attained growing importance "in many parts of the world in school, college, and university contexts" (Kunnan, 2009, p. 135). The International English Language Testing System (IELTS), Test of English as a Foreign Language (TOEFL), and Pearson Tests of English (PTE) are, for example, widely used all around the world. The increasing numbers of international students around the world and the importance of English as a means of communication in an international society have led to a considerably large population of university applicants. The scores obtained from these tests are used to make critical decisions that affect test-takers' life and prospective career. Standardized academic language proficiency test scores are frequently used for several purposes, including admissions of international students to degree programs and identification of students' post-entry language support requirements (Ockey & Gokturk, 2019). Such admission instruments are different in terms of length, format, and test content (Chalhoub-Deville & Turner, 2000) though they might be used for generically similar uses and decisions. Although TOEFL and PTE as admission tests are as valuable as IELTS, the focus of the present study is on the IELTS examination.

IELTS examination as one of the most popular tests, administered by the University of Cambridge Local Examinations Syndicate (UCLES) used for university admission purposes, predicts the extent to which a candidate will be able to begin studying through the medium of English (O'loughlin & Arkoudis, 2009). IELTS scores are required for students from particular countries to gain their visas to be admitted to English-medium universities. Through this test both academic and general English language proficiencies can be assessed.

Given the increased value placed on interactive communication in the classroom, a question arises as to whether the existing English tests for entrance to tertiary education can adequately target relevant interactional skills. Douglas (2014) argues that language use for sheer display is at best unnatural and at worst a distortion. If the performance tests elicit is in some way abnormal, the inferences made about the ability that produced the performance will stand a good chance of being wrong. Therefore, the lack of question-raising tasks in classroom interactions and language assessments raises the issue of whether the test inferences are made logically?

Gibbons (2003) suggested that students require various forms of interaction to build their four skills while direct instruction (i.e., which is characterized by highly didactic curriculum, students' receptive role, and imitation and repetition activities) remains a dominant approach in L2 education (Pufahl & Rhodes, 2011). Although language instructors express preferences for more classroom interactions, they have not always been successful at encouraging students' interactive engagement (Pianta et al., 2012). The teacher-student question-answer interaction usually deprives students of opportunities to express diverse thoughts and practice the language. Therefore, as Tan (2007) claimed, there is a vital need for more decentralized ways of teaching to promote students' conversation. Accordingly, Tocalli-Beller and Swain (2005) advocated peer-to-peer questioning in the classroom which is in line with the results of many studies suggesting that classroom interaction could be fostered through student-generated questioning (Song et al., 2017).

It follows then that if students' questioning skill is essential in interactions in or out of class, it must also be incorporated and reflected in a test that is intended to assess target language use situations (TLUS). However, IELTS applicants are rarely given a chance to reciprocally ask questions. It is assumed that their responding is parallel to questioning and in effect they are analogous. When the applicants answer questions, their questioning skill is also reflected. This assumption motivated the current research to see how these two skills compare cognitively.

2. LITERATURE REVIEW

Language assessment

The last few years, nevertheless, have seen the introduction of 'assessment' terminology into the language evaluation research discussion, signaling not merely a semantic change but a thoughtful theoretical one, with 'assessment' perceived to be a principal term used to refer to 'all methods and approaches to testing and evaluation whether in research studies or educational contexts' (Kunnan, 2004, p. 1). However, this conceptual shift goes beyond notions of alternative assessment (or alternatives in assessment, Brown & Hudson, 1998), perceiving the language evaluation process as a socially constructed activity embedded in the local context with teachers, students, and other community members recognized as meaningful assessment partners (Leung, 2004; Lynch, 2001; Lynch & Shaw, 2005; McNamara & Roever, 2006). While some available courses have reserved their language testing orientation and deal mostly with issues about the design and use of tests as the means for determining language proficiency (Bailey & Brown, 1995, and see Brown & Bailey, this issue), others have integrated additional assessment components, as well as an examination of the social roles of tests and testers in the assessment process (Kleinsasser, 2005). In language assessments, as we cannot observe the criterion directly, we use the test to make inferences about the candidates' subsequent performances. We make a distinction between the criterion which is the most relevant communicative behavior in the target situation and the test, which is often designed according to the test developer's understanding of the characteristics of the criterion (McNamara, 1997).

Speaking assessment

Assessing speaking in admission English proficiency tests revolves around the idea of whether the admitted students can cope with the language requirements of their studies. From a testing point of view, speaking is a productive skill that is interactive in nature and has to be measured directly in live interaction (Karim & Hag, 2014). In speaking tests, not only the knowledge of the language but also the ability to use it during task completion is assessed (Luoma, 2004). Therefore, linguistic knowledge areas such as grammar, vocabulary, pronunciation along with communicational, functional, and sociolinguistic knowledge are assessed as essential components in interactions. Concerning the purpose of the assessment and test interpretation, test designers decide to use different types of talks and language functions as the main focus of the assessment tasks. Various ways of speaking assessment have been proposed by many researchers (e.g. Alderson et al., 1995; Bachman & Palmer, 1996; Lynch & Davidson, 1994). Hughes (2003) suggested some standards for assessing oral ability emphasizing that to assess learners' speaking ability firstly an appropriate task to elicit information is needed and secondly validity and reliability of elicited sample and its scoring are to be ensured.

While there might be no truly authentic situation for EFL learners to practice language use and to be assessed, attempts can be made to improvise TLUS and interactions among students and their teachers (Madadi & Rezvani, 2020). Language use is related to a sociocultural view of learning as a social process framed within broader contextual practices (Vygotsky, 1986). In a Vygotskian sense, active learners do not take sole responsibility for their learning processes and for discovering

meanings. The teacher in the Vygotskian classroom carefully designs learning environments that enable learners to use languages in meaningful ways (Brevik, 2015). Likewise, a multilingual perspective on language learning considers classrooms as fundamentally social contexts in which learners use their languages as they engage in various classroom practices (Cenoz & Gorter, 2017; Garcıa & Li Wei, 2014). By doing so, students are shaped by their use of languages in communicative interaction in the social context of the classroom (Cenoz & Gorter, 2014, 2017; Park & De Costa, 2015). The interactions that learners make are part of the process of language learning and can affect their language development.

As classroom activities are based on communication and interaction, assessments must be interactive and proficiency-based as well. The learners must be put in target language use situations where they can hear and react to real uses of language. Therefore, not only they must learn to respond but also, they should learn how to pose a question to initiate or continue a communication. The appropriateness of interviews in assessing speaking and interactional competence has raised questions and concerns in language assessment (Karim & Hag, 2014) as the interactive nature of communication and more specifically TLUS calls for the need to design reciprocal tasks assessing both questions raising and responding abilities. Assessing learners through such reciprocal tasks assumes instruction on its development, though in practice teachers seldom get students to initiate or to ask questions (Graesser & Person, 1994; Rezvani & Sayyadi, 2016; Willis and Willis, 2007). Despite the instructors' remarkable resort to questioning to instruct and manage classes, they were observed to neglect and fail to encourage their learners to develop practically such a critical skill (Rezvani & Sayyadi, 2015).

Questioning in IELTS

Even though an interview as a gentle conversation between an interviewer and an interviewee is the most widely utilized task for testing speaking skills such as IELTS tests, it has been argued to have its shortcomings. In interviews, the interviewer remains dominant because he is in charge of taking all the initiatives, while the candidate or interviewee is merely supposed to respond to the questions she has been asked. Subsequently, only one style of speech is prompted, and many functions such as an inquiry for information are not characterized in the candidate's performance. According to Hughes (2003), the relationship between the interviewer and the candidate is typically such that the candidate speaks to a predominant person and he is reluctant to take the initiative. If the interviewees were given chances to pop up the questions which might cross their minds, it would, on the one hand, help them build up their confidence and be at more ease and, on the other, would aid the interviewer in assessing applicants' questioning skill (Karim & Hag, 2014).

In one of the few studies concering questioning and responding in IELTS examinations it was found that applicants who are more familiar with the format of the responding tasks and typical topics about everyday life which are commonly used in the test and practiced in IELTS preparation courses demand less information processing load (Madadi & Rezvani, 2019). To complete responding tasks the applicants had to use some ready-made chunks requiring them to use less of the cognitive mechanism resulting in short pauses and fewer hesitations (Madadi & Rezvani, 2019). To put it another way, familiar tasks imposed less information load and less demanding conditions and cognition. The question here arises as to what cognitive processes might take place when applicants face unfamiliar tasks of questioning. To this end, this paper attempts to explore whether questioning and responding as the main constituents of an authentic dialogue or discussion have analogous cognitive processes in TLUS as represented in IELTS speaking modules.

Research Question

1. How do questioning and responding in TLUS IELTS tasks compare cognitively?
2. What are the main differences between questioning and responding regarding their underlying cognitive processes?

3. METHOD

This paper described the qualitative findings of a larger project involving both qualitative and quantitative approaches. The study aimed to point out whether questioning and responding have parallel underlying cognitive processes or not. The articulated thoughts in simulated situations (ATSS) paradigm which is a think-aloud approach to cognitive assessment was used to get the participants to verbalize their thoughts while engaged in given tasks.

Participants

The sample of the study included 20 general IELTS applicants selected purposively from among applicants of two institutes in Shiraz, Iran. To see whether the proficiency of the participants might also impact the results we included both intermediate (5 male and 5 female) and advanced (5 female and 5 male) applicants. Their age ranged from 27 to 42.

Instruments

ATSS was employed to collect qualitative data in this study. The think-aloud approach is particularly useful in understanding the processes of cognition. As think-aloud methods usually assess cognitions concurrently with their occurrence, they may be better suited to tap actual thought content than other modes do (Davison, Vogel, & Coffman, 1997). In a standard think-aloud method, such as ATSS, researchers ask participants to verbalize their thoughts while performing the given tasks, and the verbalizations are recorded for subsequent analysis. The ATSS paradigm is a measure that prompts respondents' immediate cognitive and affective reactions to specific situations that just have been presented to them (Zanov & Davison, 2010). In particular, the model we used in this study is Ericsson and Simon's model (1993) known as non-metacognitive verbalization reflecting cognitive process accurately, since the applicants were not asked to explain or bring any justifications for their thoughts. As Ellis (2004) asserted this method would appear to be the most valid measure of a learner's explicit knowledge.

Concurrent think-aloud was the main method of inquiry in this study since the applicants had to think aloud while doing the tasks, though there were some unintentional reports of retrospective think-aloud as well. In this study, we were interested in thoughts the applicants had when they were completing the tasks. Often, when people are completing some tasks, they have a kind of internal monologue going through their heads, a continual stream of thoughts or feelings which mirror their reactions to something which is happening (Davison et al., 1983). The reason we opted for ATSS was that it resorts to simulated situations exactly aligned with what we intended to look at, that is, the participants' thinking in doing TLUS tasks. The tasks as ATSS prompts were questions adopted from actual general IELTS oral interviews, and respective answers representing typical questions suggested and approved by three experienced IELTS preparation course teachers. While completing the tasks, all the applicants were asked to think and talk aloud whereby their talks were recorded. We avoided interrupting the applicants but we occasionally reminded them to think aloud in case they forgot to verbalize their thoughts.

Data collection procedures

Authentic IELTS speaking tasks were given to the applicants in a counterbalancing order, half starting with questions as the first tasks and responses as the second tasks. The other group of candidates did the same tasks in reverse order to enhance the interval validity of the study. When

a sample question from the actual test was given, the applicants were asked to respond to it and when an analogous statement was prompted, they were supposed to build the original question. They were also asked to think aloud whilst they were doing the tasks. The applicants' verbal data were recorded for further analysis. The anonymity of participants' responses was emphasized to encourage open thought articulation.

Data analysis

The articulated thoughts verbalized during task completion were transcribed line by line. Then, an attempt was made to extract what the main intentions or ideas were that the participants were expressing. As a result, the researchers divided up the transcribed data into smaller parts, namely "ideas units". As Green (1998) suggested an "ideas unit" may include a single or several utterances with a single aspect of the events as the focus. The next step as suggested by Davison et al. (1983) involved grouping the ideas units into categories depending on the aim of the study and research questions. The process of identifying ideas units and categorizing was recursive and continuous as the researchers must return to raw data to re-do and re-think the transcribed data until they faced reasonable ideas unit. In accord with Glaser's (1978) recommendation every time a new idea or theme was found from the data, the researchers made notes and these memos were included in the analysis as well. To make sure whether the ideas units were consistent and reliable the inter-coder reliability of the data was examined using Cohen's kappa as an index of the agreement for categorical and sub-categorical codes. The Kappa value was 0.93 indicating substantial agreement for the coding (see Landis & Koch, 1977 for Kappa index interpretations).

4. RESULTS AND DISCUSSION

The first set of associated questions and answers concerned the participants' favorite food and TV programs; ten participants answered an actual IELTS question (What's your favorite flavor?) and the rest tried to build a question for a typical answer to it (My favorite flavor is sweet.). All the applicants had a clear expression of their opinion although they had some pauses (silent and filled) thinking about what to answer. The given question seemed rather challenging since most of the applicants didn't know what their real flavors were. A point worthy of note in most verbal reports was that they first identified the keywords of the question. This was followed by thoughts of how it could be structured in the answer. Having identified the keywords semantically or grammatically as the first step, they moved on to structure them into simple statements as the next step. Then, challenges aroused as most of them could not decide what their favorite flavors were. As the following example suggested while one of the intermediate applicants (extract A) recognized the appropriate structures to answer the question he had a silent pause since he was not sure about his favorite flavor. The given response by the applicant manifested that the pause couldn't be because of a lack of grammatical knowledge as the answer was grammatically and semantically correct, thus the applicant had doubts about the flavor. This might be admitted as the applicant had used the word "think" which is a sign of doubt.

- Extract A (intermediate)

Question: What is your favorite flavor?

Responding Verbatim think-aloud: I (1) think spicy is my favorite flavor.

*numbers in parentheses indicate the pause length.

The other advanced applicants (B & C) answered the same question with a filled pause (um), though the pause wasn't suggestive of a lack of grammar or lexical knowledge since the applicants responded accurately and used more lexis to complete the question in longer statements. It was observed that filled pauses were used by advanced applicants in comparison to intermediate applicants, who on the contrary had more tendency to pause silently.

- Extract B (advanced)

Question: What is your favorite flavor?

Responding Verbatim think-aloud: I like (um,1) mint and lemon.

- Extract C (advanced)

Question: What is your favorite flavor?

Responding verbatim think-aloud: I prefer (um,1) chocolates and candies (1) which their flavor is sweet.

Based on the data analysis "pause length" was identified as the first "ideas unit". Completing the responding task, all the applicants seemed to reduce the length of pauses while giving the responses. The applicants' short pauses were natural as native speakers also might normally pause while answering such questions. Consequently, effective pauses not only were beneficial for comprehension but also, were the manifestation of more natural and native-like speech. This is analogous with the finding that pausing is natural and necessary for breathing needs and for pragmatic use during speech (Ling, 2006), though long and unnecessary pauses might have negative effects on speaking fluency.

Fluency was considered as the second "ideas unit". Levelt (1989) emphasized that speaking involves the processing of a considerable amount of data in a limited period, that is, two or three words are produced per second in natural speech. He claimed that fulfilling this great task requires automaticity, not conscious monitoring as human capacity is too limited to focus consciously on the information (Segalowitz & Hulstijn, 2005). Other researchers also claimed that a pivotal difference between fluent and non-fluent L2 speakers is the extent to which lexical processing is automatized. Different factors such as speech rate, repairs, amount and frequency of hesitation, location of pauses, and length of runs of fluent speech between pauses are associated with the psycholinguistic facets of performance and production (e.g. Lennon, 1990; Möhle, 1984; Towell et al., 1996). Levelt (1989) also maintained that the automaticity of language production has generated uninterrupted fluent speech, therefore short pauses in the responding tasks in addition to the correct form of responses might lead to the fluency of the applicants in responding tasks. As such, and as fluency and connected speech are interrelated, it can be inferred from the applicants' verbal reports that both intermediate and advanced applicants had used more connected speech in their given responses and were considered as more fluent.

This might suggest that information load or cognitive processing demand (Krivokapić, 2007) is an effective factor regarding the length of pauses. Information load is likely to be minimal if the learners are familiar with the tasks they have to perform. It would also be of a minimum load if the language learners are familiar with the topics they are going to talk about. Higher information load, in turn, makes cognition more careful and slower (Rabbitt, 1968; Robinson, 2001). This is in agreement with other studies concluding that learners that were more familiar with the format of the responding tasks and typical topics about everyday life which are commonly used in the test and practiced in IELTS preparation courses very often demand little information processing load (Madadi & Rezvani, 2019).

In questioning tasks on the other hand (extract D) borderlines between lexical words were more clearly identified through long pauses resulting in less natural speaking. To complete questioning tasks the applicants had to construct questions from scratch, requiring them to use more of the cognitive mechanism resulting in a longer pause as an indicator of cognitive processing (Madadi & Rezvani, 2019). In other words, unfamiliar tasks imposed extra information load and more demanding conditions and cognition. The long pauses would break up the questions into smaller parts leading to less connected and natural speech.

- Extract D (advanced):

Response: Yes, I think holidays are becoming more and more important.

Questioning verbatim think-aloud (advanced): (Uh,3) Do you think (2) we should have a (2) program (1) or plan for (2) on holidays, (1) for our holidays

It was observed that the pause length in these tasks increased to 4-6 seconds and self-correction strategies were used a lot to construct or recast an appropriate question as manifested in extract E and F.

- Extract E (advanced):

Response: I want to travel to the places which are unique and interesting.

Questioning verbatim think-aloud (advanced): (Uh,4) which/ what kind of places do you want to go?

- Extract F (intermediate):

Response: I prefer traveling in a group rather than traveling alone.

Questioning verbatim think-aloud (advanced): (Uh,3) Do you prefer to go to trip (1) with ($\rightarrow$ with omitted) alone or (2) with somebody/ with someone.

The analysis and comparison of the applicants' verbal reports revealed that IELTS applicants tended to be more fluent and native-like in responding tasks than questioning tasks irrespective of their proficiency levels. Similarly, it was found that connected speech resulting in more fluency contributed toward naturalness and intelligibility. Intelligibility has to do with 'speech clarity' or the proportion of a speaker's output that a listener can readily understand. Hence, more fluent applicants were more intelligible as well.

Focusing on verbal reports and transcriptions, we inferred that the applicants had good knowledge of grammar, as far as the responding tasks were concerned, since all the applicants regardless of their proficiency levels were able to answer the questions correctly. More specifically, their verbal reports also manifested that in addition to knowing the grammatical rules theoretically, the applicants were also able to operationalize the rules to make correct answers to the questions. So, both explicit (verbal reports) and implicit (correct answers to the questions) knowledge of grammar were available among all the applicants while doing the responding tasks. As a result, "knowledge of grammar" was identified as the "ideas unit" closely associated with accuracy. The following exemplary excerpts evidence the findings.

- Extract G (intermediate):

Question: What's your favorite flavor?

Responding Verbatim think-aloud 1 (intermediate): "this is a WH question [!]", "so I should initiate the answer like this" "my favorite flavor is (1) let's say dark, my favorite flavor is dark."

Responding Verbatim think-aloud 2 (advanced): (1), "a WH question, with favorite as the focus", ok. "My favorite flavor is lemon and mint."

In questioning tasks, the same ideas unit was identified in their verbal reports, though the applicants could not accomplish accuracy. As the following sample extracts exemplify, the applicants of both levels (intermediate: H & I, advanced: J & K) tried to focus on using grammatically correct structures to make related questions, but they were not quite successful. The applicants centered their attention on the correct forms in making questions as the recorded verbal reports of the applicants' articulated thoughts exhibit.

- Extract H (intermediate):

Response: I may visit my family or have a short trip in holiday.

Questioning Verbatim think-aloud: ok, I should use "do", (1) do you may visit your family? Or let's say it without "May", (2) do you visit in holiday?

- Extract I (intermediate)

Response: I want to travel to places which are unique and interesting.

Questioning Verbatim think-aloud: where, (2) (should begin with where and want). Where do you want, no, no, no, like to travel to? Where do you like (1) to travel in future?

- Extract J (advanced)

Response: I may visit my family or have a short trip in holidays.

Questioning Verbatim think-aloud 1: (1) "you want me to make a question?" (2), "I should use part of the sentence as a part of question," (uh, 2), about holidays (2), "which kind of holiday you visited your parents?"

Questioning Verbatim think-aloud 2: a WH question about holidays/ (3) it might be a general question (um, 2) When did you (1) when did you visit your family?

- Extract K (advanced)

Response: Yes, I think holidays are becoming more and more important?

Questioning Verbatim think-aloud: "It needs a yes/no question, (um, 2) which holidays do you know", (2) "and you prefer it's important for you?"

In the sample extracts (H & I) the intermediate applicants were verbalizing the correct structural rules to accomplish the questioning tasks, though they couldn't use them to build up questions. Other sample extracts (J & K) also pointed to the same problem among advanced applicants. To conclude, most applicants irrespective of their proficiency level couldn't apply the verbalized and well-known rules of structures while performing the questioning tasks.

Another point worthy of note is the applicants' conscious and unconscious thoughts and strategies to smooth task completion (for a discussion see Cohen, Ledford, & Spreitzer, 1996) and to personalize the task accomplishment procedure. In both tasks completion the applications of planning, monitoring, and evaluating as metacognitive strategies can be identified. Metacognition refers to awareness of one's knowledge, what one does and does not know, and one's ability to understand, control, and manipulate one's cognitive processes (Meichenbaum, 1985). It includes knowing when and where to use particular strategies for learning and problem solving as well as how and why to use specific strategies. Chamot (2005) maintains that strategic language learners hold metacognitive knowledge about their thinking and learning approaches, a good understanding of what a task entails, and the ability to organize the strategies that best suit both the task demands and their strengths. The language learners do not think about these strategies while performing them but, if they are asked what they are doing, they can usually accurately describe their metacognitive processes and strategies, therefore the applicants of this study were asked to think and talk aloud while completing the tasks to pinpoint what metacognitive strategies they had used during task completion. As an example, the articulated thoughts of an applicant (extract J) are provided below. it can be seen that the applicant was thinking about preparation for task execution, identifying procedures and requirements of a task to plan it, by asking the question that "you want me to make a question?"

- Extract J (advanced)

Response: I may visit my family or have a short trip in holidays.

Questioning Verbatim think-aloud 1: (1) "you want me to make a question?" (2), "I should use part of the sentence as a part of question," (uh, 2), about holidays (2), "which kind of holiday you visited your parents?"

There were also generated thoughts that represented the recognition of a problem (grammatical errors, inappropriate lexical choice) followed by no further attempt to correct it (extract L) or evaluations that lead to improvements or revisions through self-correction (extract M). These examples point to the monitoring and evaluation processes. Such examples of strategies use led to the identification of the metacognitive strategies as another "ideas unit".

- Extract L (intermediate)

Response: I prefer traveling in a group rather than traveling alone.

Questioning Verbatim think-aloud: (4) how do you like travel? I think, I'm not sure. Something is missing.

- Extract M (advanced)

Response: I prefer traveling in a group rather than traveling alone.

Questioning Verbatim think-aloud: (Uh, 3) Who (1) do you prefer to go on a trip? (1) On a trip with (self-correction)?

In both tasks, metacognitive strategies were frequently used by the applicants. However, in responding tasks, the operation of the metacognitive strategies such as self-correction yielded almost completely correct responses, while in questioning tasks the applicants did evaluate the questions, but very often they failed to self-correct even if they could identify the problem. The following is a telling example that shows that the applicant identified the problem, self-corrected but the cognitive demands and loads disrupted the memory of the questioning task.

- Extract N (advanced)

Response: Yes, I think holidays are becoming more and more important.

Questioning verbatim think-aloud: (Uh, 3) do you think we should have a (2) program or plan for on holidays (self-correction), (1) for our holidays? Would you please repeat the sentence?

Overall, we found more errors in questioning tasks though the lexis in both questions and responses produced were of a similar size. The errors for measuring accuracy included lexical, morphological, and syntactic errors and omissions of article, verb, and subject (see Michel, Kuiken, & Vedder, 2007).

It was also noted that the applicants used more ready-made chunks to respond to questions. This also reduced the pause length and the number of errors. It can be inferred that as the participants had practiced similar tasks they remembered or were coached to provide quicker and more formulaic responses. This also minimizes the cognitive load and frees more of the cognitive capacity for analysis as it is argued that in producing routines speakers do not necessarily implement all stages of speech production (Pickering & Garrod, 2004). Kormos (2011), comparing routines and creatively constructed elements, also contended that routines follow the same processing stages faster, and with less conscious effort, resulting in shorter pauses and less cognitive load. When one needs to create novel plans or creative responses, consciousness is necessary for short- and long-term planning (Mandler, 1975).

In view of speech production, the control of the articulatory system might be unconscious but planning what to say might be conscious, particularly if one is expressing some new ideas, or expressing some old ideas in a novel way. According to Bock (1982), syntactic planning by skilled speakers is also relatively automatic and outside conscious voluntary control. This provides an account of why the IELTS applicants known as skilled speakers, had rather more automatic

syntactic planning or unconscious/subconscious process leading to faster-responding task completion because of the greater role of task familiarity and ready-made chunks. In the following task, the applicant could easily employ a recurrent and familiar statement.

- Extract O

Question: What is your favorite flavor?

Response verbatim think-aloud 1 (intermediate): my favorite flavor is sweet.

Response verbatim think-aloud 1 (advanced): my favorite flavor is dark, (1) like dark chocolate.

In question construction tasks, on the other hand, a completely novel task must be done. Planning what to say, finding relevant lexis, and putting them into meaningful linguistic forms, however, require more cognitive efforts. Jackson (1982) suggested that the amount of planning required depends on whether the speech is "new" speech or "old" speech. Old speech (well-known phrases) requires little planning and is relatively continuous. New speech demands planning and is characterized by hesitation pauses as indicated in this study. In Skehan's words (1998) and in short, cognitive familiarity and cognitive processing are two effective factors concerning cognitive complexity. Madadi and Rezvani (2019) further argued that applicants' unfamiliarity with questioning as an untrodden path augments the task difficulty and results in more and longer pauses, and turn, in less natural/ native-like speech.

To sum up, the following ideas units were identified as the researchers organized the transcriptions for easy retrieval. The pause and length of the pauses, the applicant's ability to operationalize their knowledge of grammar, metacognitive strategies such as self-evaluation and self-correction, and the use of readymade linguistic chunks were categorized as "ideas units".

Under the emerging codes, the following results have been noted and discussed. Table 2 summarizes the findings

Table 1: Ideas Units from Stimulated Recall Protocols

1. Pause length
2. fluency
3. Knowledge of grammar
4. Metacognitive strategies
5. Ready-made chunks

Table 2: Main Differences between Questioning and Responding Tasks

questioning	responding
less accurate	more accurate
less fluent	more fluent
more pauses/longer pauses	fewer pause/shorter pauses
less connected speech	more connected speech
less use of metacognitive strategies	more use of metacognitive strategies
less natural	more natural

5. CONCLUSION AND IMPLICATIONS

The overall findings emerging from the qualitative analysis of the verbal data indicated that responding tasks were considered as familiar tasks leading the applicants to more accurate and fluent language production, hence the applicants had fewer grammatical errors while completing the responding tasks. More self-corrections were found in responding tasks by using metacognitive strategies. More use was also made of routine linguistic chunks in responding tasks, while in questioning tasks, the applicants attempted to put more novel words together making the task, even more demanding with longer pauses. The findings generally support Plato's account of Socrates stating that, contrary to the general opinion, it is more difficult to ask questions than to answer. From a psycholinguistic point of view, questioning and responding though assumed to be parallel, differ to a noticeable extent in terms of the underlying cognitive processes. The more noticeable accuracy and fluency in responding can be linked to the applicants' familiarity with the frequent responding role students take in education. This reduces the cognitive load and efforts making the differences.

The results of the study might have implications for second language learners, teachers, and test developers. Tests can have effects on learners' and teacher's behaviors in the classroom, and "impact" more widely on teaching materials, educational systems, and even society (Taylor, 2005; Weir & Milanovic, 2003). Indeed, washback can be considered to be one aspect of impact (Taylor, 2005), the former known as micro-level and the latter as macro-level effects (Weir & Milanovic, 2003). The influence of tests on students' learning and future is irrefutable given particularly the high stakes of IELTS examinations. As the study demonstrated there are cognitive differences between questioning and responding underlying processes that merit attention and inclusion in language assessment. The current responding focus, for learners indeed, induces them to hence focus on the development of responding strategies and linguistics forms and lexis.

As Alderson and Wall (1993) argue examinations are expected to influence teachers' content of teaching and how they teach or in the words of Cheng et al. (2015) teachers play an important catalyst role in carrying positive and negative washback effects of tests. Thus, it is advisable to give a more active part to the students and their questioning in L2 classes. We think that this will pay off and active participation and question making will keep or generate new communications, essential in L2 learning.

Negative washback occurs in situations where there may be a discrepancy between the desired goals of instruction and the focus of assessment. An implication of the results of this study is that test developers, classroom or high stakes, should include space for the test-takers initiation of dialogues and questions by incorporating more reciprocal tasks instead of only requiring them to rather passively answer teachers' or examiners' questions. This can lead to a more positive washback for both L2 learners and teachers as discussed above.

References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing.* American Educational Research Association.

Alderson, C. J., Alderson, J. C., Clapham, C., Wall, D., & Swan, M. (1995). *Language test construction and evaluation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511732980

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129. https://doi.org/10.1093/applin/14.2.115.

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

Brown, J. D., & Bailey, K. M. (1995). Language testing courses: What are they in 2007? *Language Testing, 25*(3), 349-383. https://doi.org/10.1177/0265532208090157

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*(1), 1-30. https://doi.org/10.1037/0033-295X.89.1.1

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*(4), 653-675. https://doi.org/10.2307/3587999

Cenoz, J., & Gorter, D. (2014). *Focus on multilingualism as an approach in educational contexts.* Springer. https://doi.org/10.1007/978-3-319-02261-6

Cenoz, J., & Gorter, D. (2017). Minority languages and sustainable translanguaging: Threat or opportunity? *Journal of Multilingual and Multicultural Development, 38*(10), 901-912. https://doi.org/10.1080/01434632.2016.1263162.

Chalhoub-Deville, M., & Turner, C. E. (2000). What to look for in ESL admission tests: Cambridge certificate exams, IELTS, and TOEFL. *System, 28*(4), 523-539. https://doi.org/10.1016/S0346-251X(00)00033-5

Chamot, A. U. (2005). Language learning strategy instruction: Current issues and research. *Annual Review of Applied Linguistics, 25*(1), 112-130. https://doi.org/10.1017/S0267190505000061

Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching, 48*(4), 436-470. https://doi.org/10.1017/S0261444815000233.

Cohen, S. G., Ledford Jr, G. E., & Spreitzer, G. M. (1996). A predictive model of self-managing work team effectiveness. *Human Relations, 49*(5), 643-676. https://doi.org/10.1177/001872679604900506

Davison, G. C., Robins, C., & Johnson, M. K. (1983). Articulated thoughts during simulated situations: A paradigm for studying cognition in emotion and behavior. *Cognitive Therapy and Research, 7*(1), 17-39. https://doi.org/10.1007/BF01173421

Davison, G. C., Vogel, R. S., & Coffman, S. G. (1997). Think-aloud approaches to cognitive assessment and the articulated thoughts in simulated situations paradigm. *Journal of Consulting and Clinical Psychology, 65*(6), 950-958. https://doi.org/10.1037/0022-006X.65.6.950.

Douglas, D. (2014). *Understanding language testing.* Routledge. https://doi.org/10.4324/9781315889771

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning, 54*(2), 227-275. https://doi.org/10.1111/j.1467-9922.2004.00255.x.

Ericsson, K. A., & Simon, H. A. (Eds.). (1993). *Protocol analysis: Verbal reports as data*. Cambridge University Press.

Fox, J. (2008). Alternative assessment. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of language and education. Language testing and assessment* (pp. 97-109). Springer. https://doi.org/10.1007/978-3-319-02261-6_8

García, O., & Wei, L. (2014). Language, bilingualism, and education. In *Translanguaging: Language, bilingualism, and education* (pp. 46-62). Palgrave Pivot. https://doi.org/10.1057/9781137385765

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly, 37*(2), 247-273. https://doi.org/10.2307/3588502

Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. University of California.

Graesser, A. C., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology, 10*(7), 17-31. https://doi.org/10.1002/(SICI)1099-0720(199603)10:1<17::AID-ACP368>3.0.CO;2-V

Graesser, A., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal, 31*(1), 104-137. https://doi.org/10.3102/00028312031001104

Green, A. (1998). *Verbal protocol analysis in language testing research: A handbook*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667329

Hughes, A. (2003). *Testing for language teachers*. Cambridge University Press.

Jackson, G. A. (1982). Public efficiency and private choice in higher education. *Educational Evaluation and Policy Analysis, 4*(2), 237-247. https://doi.org/10.3102/01623737004002237

Karim, S., & Haq, N. (2014). An assessment of IELTS speaking test. *International Journal of Evaluation and Research in Education, 3*(3), 152-157.

Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect, 20*(1), 77-102.

Kormos, J. (2011). Speech production and the Cognition Hypothesis. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 39-60). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.2.04kor

Krivokapić, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics, 35*(2), 162-179. https://doi.org/10.1016/j.wocn.2006.01.002

Kunnan, A. J. (2004). Regarding language assessment. *Language Assessment Quarterly, 1*(1), 1-4. https://doi.org/10.1207/s15434311laq0101_1

Kunnan, A. J. (2009). Testing for citizenship: The US naturalization test. *Language Assessment Quarterly, 6*(1), 89-97. https://doi.org/10.1080/15434300802606897.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 33(1), 363-374. https://doi.org/10.2307/2529786

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387-417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x

Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. *Language Assessment Quarterly, 1*(1), 19-41. https://doi.org/10.1207/s15434311laq0101_3

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524605.

Ling, H. (2006). Long pauses in Chinese EFL learners' speech production. *Interlingüística, 17*(1), 606-616.

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.

Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing, 18*(4), 351-372. https://doi.org/10.1177/026553220101800402

Lynch, B., & Shaw, P. (2005). Portfolios, power, and ethics. *TESOL Quarterly, 39*(2), 263-297. https://doi.org/10.2307/3588311

Lynch, B. K., & Davidson, F. (1994). Criterion-referenced language test development: Linking curricula, teachers, and tests. *TESOL Quarterly, 28*(4), 727-743. https://doi.org/10.2307/3587553

Madadi, M., & Rezvani, R. (2019). Trodden and untrodden paths: A study of cognitive processes in oral responding and questioning. *Journal of Teaching Language Skills, 38*(3), 159-187.

Mandler, J. M., & Day, J. (1975). Memory for orientation of forms as a function of their meaningfulness and complexity. *Journal of Experimental Child Psychology, 20*(3), 430-443. https://doi.org/10.1016/0022-0965(75)90088-6

McNamara, T. F. (1997). Interaction in second language performance assessment: Whose performance? *Applied Linguistics, 18*(4), 446-466. https://doi.org/10.1093/applin/18.4.446.

McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Oxford University Press.

Meichenbaum, D. (1985). Teaching thinking: A cognitive-behavioral perspective. In J. W. Segal, S. F. Chipman, & R. Glaser (Eds.), *Thinking and learning skills* (Vol. 2, pp. 407-426). Routledge.

Michel, M., Kuiken, F., & Vedder, I. (2007). The interaction of task condition and task complexity in the oral performance of Turkish and Moroccan learners of Dutch. In *2nd International Conference on Task-based Language Teaching, Hawaii, USA*. https://doi.org/10.1075/tblt.2.09mic

Moehle, J. P. (1984). Seismic response of vertically irregular structures. *Journal of Structural Engineering, 110*(9), 2002-2014. https://doi.org/10.1061/(ASCE)0733-9445(1984)110:9(2002)(

Ockey, G. J., & Gokturk, N. (2019). Standardized language proficiency tests in higher education. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 1-16). Springer. https://doi.org/10.1007/978-3-030-02899-2.

O'Loughlin, K., & Arkoudis, S. (2009). Investigating IELTS exit score gains in higher education. *International English Language Testing System (IELTS), 10*(1), 1-86.

Park, J. H., & De Costa, P. (2015). Reframing graduate student writing strategies from an activity theory perspective. *Language and Sociocultural Theory, 2*(1), 25-50. https://doi.org/10.1558/lst.v2i1.26546

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365-386). Springer. https://doi.org/10.1007/978-1-4614-2018-7_17

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169-190. https://doi.org/10.1017/S0140525X04000056

Pufahl, I., & Rhodes, N. C. (2011). Foreign language instruction in US schools: Results of a national survey of elementary and secondary schools. *Foreign Language Annals, 44*(2), 258-288. https://doi.org/10.1111/j.1944-9720.2011.01130.x.

Rabbitt, P. M. (1968). Channel-capacity, intelligibility, and immediate memory. *The Quarterly Journal of Experimental Psychology, 20*(3), 241-248. https://doi.org/10.1080/14640746808400158

Read, J. (2009). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes, 8*(1), 180-190. https://doi.org/10.1016/j.jeap.2008.10.002.

Rezvani, R., & Sayyadi, A. (2015). Instructors and learners' questioning: A case of EFL classroom discourse in Iran. *Journal of Teaching Language Skills, 34*(3), 141-164.

Rezvani, R., & Sayyadi, A. (2016). Washback effects of the new Iranian TEFL Ph.D. program entrance exam on EFL instructors' teaching methodology, class assessment, and syllabus design: A qualitative scrutiny. *Journal of Instruction and Evaluation, 9*(33), 155-176.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge University Press. https://doi.org/10.1017/CBO9781139524780.012

Seedhouse, P., & Egbert, M. (2006). The interactional organization of the IELTS speaking test. *International English Language Testing System (IELTS), 6*(1), 35-52. https://doi.org/10.1016/j.specom.2007.12.012.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). Oxford University Press.

Simon, M., Ercikan, K., & Rousseau, M. (Eds.). (2012). *Improving large-scale assessment in education: Theory, issues, and practice*. Routledge. https://doi.org/10.4324/9780203801006

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press. https://doi.org/10.1017/CBO9780511734953

Song, D., Oh, E. Y., & Glazewski, K. (2017). Student-generated questioning activity in second language courses using a customized personal response system: A case study. *Educational Technology Research and Development, 65*(6), 1425-1449. https://doi.org/10.1007/s11423-017-9536-1

Tan, Z. (2007). Questioning in Chinese university EL classrooms: What lies beyond it? *RELC Journal, 38*(1), 87-103. https://doi.org/10.1177/0033688206076169

Taylor, L. (2005). Washback and impact. *ELT Journal, 59*(2), 154-155. https://doi.org/10.1093/eltj/cci030

Tocalli-Beller, A., & Swain, M. (2005). Reformulation: The cognitive conflict and L2 learning it generates. *International Journal of Applied Linguistics, 15*(1), 5-28. https://doi.org/10.1111/j.1473-4192.2005.00079.x

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84-119. https://doi.org/10.1093/applin/17.1.84.

Weir, C. J., & Milanovic, M. (Eds.). (2003). *Continuity and innovation: Revising the Cambridge Proficiency in English Examination*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667398

Wendler, C. L., & Walker, M. E. (2006). Practical issues in designing and maintaining multiple test forms for large-scale programs. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 445-467). Routledge. https://doi.org/10.4324/9780203874772

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.

Zanov, M. V., & Davison, G. C. (2010). A conceptual and empirical review of 25 years of cognitive assessment using the articulated thoughts in simulated situations (ATSS) think-aloud paradigm. *Cognitive Therapy and Research, 34*(3), 282-291. https://doi.org/10.1007/s10608-009-9255-9

TTC for TEYL: From Assumed Competence to Observed Performance

Saeedeh Shamsaee[1,*] iD & Mohammad Ali Shams[2]

[1] Mazandaran University of Science and Technology, Babol, Iran
[2] Babol University of Medical Sciences, Babol, Iran

Abstract
The present study aimed to educate a group of young teachers through a TTC for TEYL while evaluating their assumed competence and observed performance. Fifteen student teachers participated in a 5-session online course and maintained their involvement through offline group discussions between weekly sessions. After the course, they were given a 25-item questionnaire to express their subjective assessment in regard with the effect of the TTC on their teaching competence, and were also required to develop multimedia TEYL materials reflecting their actual grasp of the presented content. A chi square test showed that the answers given to 11 items of the questionnaire across five dimensions of TEYL namely motivation, mediation, maturation, manipulation and modulation produced significant coefficient reflecting the teachers' perception of their abilities. In addition, two independent raters scored their performance on the task. The assumed competence and the observed performance were quantified and analyzed in a linear regression model to see if the former can predict the latter. Moreover, the individual and collective reactions of the student teachers both to the contents of the online course and issues raised in the offline discussions were examined via educational discourse analysis. The findings set the ground for designing and assessing future TTC courses in TEYL.

1. INTRODUCTION

Teacher education in Iran is still young, and when it comes to teaching English to young learners (TEYL), the need for established and effective teacher training courses (TTC) is felt the most. This is rather consistent with the global realization with regard to the dire need for research-based benchmarks in this area. A wide range of evaluation systems might be presented for assessing the efficiency of TTCs for TEYL; however, one has to be able to select among the frameworks based on the objectives of evaluation. In addition, the nature and mechanisms of TTCs are variously designed according to the factors contributing to the educational markets available in each country. We first describe course at hand, and then elaborate on its empirical counterparts along with theoretical rationales only when it is really needed for the sake of operational accuracy and clarity.

* **Corresponding Author:** Saeedeh Shamsaee, English Department, Mazandaran University of Science and Technology, Babol, Iran. **Email**: saeedeh.shamsaee@mustmb.ac.ir
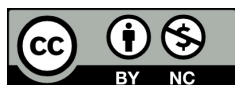
It is worth mentioning that this approach is inspired by action research agendas which have recently been celebrated by both educational researchers and teacher educators more than ever.

In action research, practice is not confined by theory; first practical problems are dealt with based on experience and then a research framework is adopted to explore and evaluate the extent of failure or success achieved in practice. The researcher's experience in TEYL and Teacher Education was organized in a TTC, and the present study was conducted to evaluate it. Specifically we were interested to examine both subjective perspectives and objective output associated with the course; therefore, we first educated the student teachers, then examined their opinion about the effectiveness of the course and finally scored their performance by examining the teaching material produced by them at the end of the course. We designed an online 5-session workshop for the student teachers who were interested in embarking on a TEYL career. The online space, which at first might seem a liability imposed by the COVID-19 pandemic, proved to be a blessing in disguise, and helped us come up with innovative lesson plans and creative assignments. Due to a general negligence toward the significance of TEYL itself, the currently available TTCs in this area are even further deprived of updated contributions and field-associated improvements. The five sessions were allocated to five dimensions of teacher education identified in our literature review. To examine the effect of the TTC on the student teacher's perception of their abilities and also its relationship with their actual teaching performance, the research questions were formulated follows:

1) How do student teachers react to the five dimensions of TCC for TEYL presented during the training course?
2) What do they assume (while answering a questionnaire) about the effect of the course on their teaching competence?
3) Does their assumed competence meaningfully predict their observed performance on a material development task?

The following five sections elaborate on the five dimensions and their background in applied linguistics. As mentioned above, we adhere to the principle of operational clarity, and avoid being entangled in a web of vague theoretical citations.

2. REVIEW OF THE LITERATURE

Motivation in TEYL

There are two main dimensions for teaching motivation in TEYL: willingness and confidence. However, these two concepts may appear in various wordings. Strakova (2015) who worked on the self-efficacy of pre-service trainees uses the term "readiness". Sad (2015) labels these two dimensions as perception of efficacy and willingness to teach. His study shows that almost half of the variation observed in perceived efficacy can be explained by teachers' perception of their language proficiency. In other words, the more proficient the teacher assume they are, the higher is the chance of being confident about their capability in teaching English to young learners. In another project, Camlibel-Acar (2017) asked a group of young teachers about their perceived capability and readiness to teach English to young learners before and after a TTC, and found that although the quantity of the responses had not changed significantly, the quality of their arguments had definitely increased. Dealing with immature learners who are totally immersed in their newly discovered emotions is painfully demanding. In the absence of an intrinsic motivation, one needs a very powerful extrinsic reason to accept such a challenge. Our first session was set to disillusion student teachers and help them face the reality of the issue at hand. Although the challenges imposed on ELT practitioners because of the expansion of TEYL have been discussed for at least two decades (see Cameron, 2003), most student teachers who embark on TEYL are under the

wrong impression that it is easier to teach children because their language needs are much easier to handle, and even the least level of proficiency and teaching aptitude on the part of the teacher would suffice to meet their needs.

This oversimplification contributes to a false teaching confidence that vanquishes at their first real teaching experience. Copland, Garton and Burns (2014) provide a rather comprehensive review of the global challenges faced by TEYL practitioners, and call for a more realistic approach. Therefore, aiming to help them build a rather realistic confidence, we first informed the student teachers with a combination of empirical evidence and real-life examples that exhibited the main difficulties associated with language classrooms for children. This might seem quite counterintuitive since to motivate, one usually uses the positive side of things; however, in our approach disillusionment was considered an indispensable part of the foundation needed for establishing an enduring motivation. In Woolfolk's (2008) words, defines self-efficacy is a "teacher's belief that he or she can reach even difficult students to help them learn" (p. 361); however, if the teacher is mistaken about the meaning and instance of that "difficulty", their assumption of self-efficacy will be false. Therefore, if a young teacher maintains that they are confident and willing to teach children, but they are unaware of the real difficulties of TEYL, the confidence is false, and the willingness is soon to be vanished. In addition, perceptions need to be constantly updated. Concepts such as teaching motivation and self-efficacy cannot remain intact when the whole world is involved in a pandemic. For this very reason, a researcher such as Dorsah (2021) tries to redefine difficulties of teaching and reconstruct the concept of "teaching readiness" when remote learning has become a must rather than a fancy option. In accordance with this view we assumed that real pedagogical affection is based on true knowledge and realistic vision as opposed to transitory fantasies about the teaching profession. The current research has focused on the extent of the effect exerted by our treatment (training course) from the perspective of the teachers.

Mediation in TEYL

Medium-awareness plays a vitally important role when teaching English to children. As Ratminingsihm, Mahadewi and Divayana (2018) emphasize, the success of any TEYL program is closely associated with the teachers' grasp of the most recent information and communication technologies. In their ICT-based approach to TEYL, the teachers were educated to be updated with regard to progressive platforms and fully equipped mediating systems played a central role; the results of their TTC study show that such an approach has a significant effect on the young learners' enthusiasm and concentration during learning hours. The recognition of the determining place of E-learning in TEYL is not a recent development; around twenty years ago, Chujo and Nishigaki (2004) tended to the issue and designed an interactive TTC workshop in which student teachers could discuss and come up with their individualized online frameworks. Teacher educators who train young teachers for teaching English to young learners cannot ignore the crucial importance of emerging technologies; as Pin (2013) puts it, a balanced use of such technologies is a must for teachers who have to deal with young learners' emerging minds; this is now even more obvious given the global necessities for remote learning. As mentioned earlier in the section on motivation, in a very recent study, Dorsah (2021) examined pre-service teachers' readiness for remote learning in the wake of COVID-19. This study was important to us both for its methodological integrity and realistic approach to the issue of mediation in teaching. Dorsah's emphasis on online leaning readiness and emergency remote teaching are well in line with the mediation dimension in our TTC and the questionnaire that sought to assess its effect on the teachers' competence. The idea was to make the young teachers aware of the dire need for a mastery of online mediating systems and devices; their reaction will be discussed in our results.

Maturation in TEYL

Our operational definition of maturation can be explained in three areas namely attention, connection and parenting. Teaching maturity in TEYL, as we defined it in the present study, refers to the teachers' ability in examining children' learning behavior, communicating with them and playing the role of contemporary parent during the learning hours. The last component is also associated with the issue of parent involvement since building a parent-like relationship with young learners requires help from the parents themselves. The teachers' life history (micro level), and their relationship with other factors in the relevant educational context including parents (meso level) is crucially important in TEYL (Sowa, 2017). She explains how the interaction of these two levels forms teacher's identity as one of the caregivers in the eyes of the children. In TEYL, teacher becomes the third side of a triangular relationship which had been previously managed only by children and their parents. The above- mentioned maturation is in fact the teacher's efficacy in becoming an effective third party in the triangle.

It is worth mentioning that in some studies, the parent-like role of the successful TEYL teachers appears with alternative wordings, but the nature of the description directs us to the same aspects discussed under the maturation dimension in the present study. For example, in the findings of the study done by Hussein (2014) on the characteristics of exemplary TEYL teachers, the personal features include the following items: inspiring, expressing true love and sincerity, kind, friendly, patient, caring, enthusiastic, attractive, motivating, helpful, creative and communicative. It is hard not to see that many of the above qualities overlap with those of desired parent. The communicative aspect of the maturation dimension is mainly visible in the language used by teachers. Ahmad and Samad (2018) did an innovative study on the proxy-like functions of metaphoric expressions in the speech of TEYL teachers and explained how these communicative components contribute to the formation of teacher's role as a semi-parent who uses parent-like expressions in their communication with young learners. The participants' reaction to the maturation dimension in our TTC will be discussed in our results.

Manipulation in TEYL

Manipulation here refers to teachers' ability in owning the teaching material through planning, tailoring and development. A lack or deficiency of TEYL teachers' manipulative skills would lead to the mechanical use of irrelevant material or the unauthenticated use of standard material. The concept of authentication as a teaching process as opposed to authenticity as a static feature attached only to certain teaching materials was introduced by Nunan (1997). He maintained that beside the authenticity of the materials, learners should be able to relate the material to their own background. Well in case of young learners which have not gained enough autonomy yet, the main burden is on the teacher. Cao (2019) examined the quality of TEYL curriculums in Vietnam and emphasized that materials, class atmosphere and other contextual factors should all work together. In other words, the process of authentication should accompany the authentic products used as learning materials. Rich (2014), in an overview of TEYL research and practice, explained the central role that realistic materials can play in keeping pace with the realities of children's life.

Modulation in TEYL

Modulation in our terms is concerned with language skills and language components in TEYL. During the training course, the student teachers were provided with an insight into the place and priority of each module for young learners. The modulation section in our questionnaire examined the participants' perception of the TTC material across the seven modules. Different groups of teachers, based on their background, might have different priorities in mind when it comes to language modules in TEYL. Wissink and Starks (2019) developed a questionnaire to explore the

teachers' demands for a TTC in TEYL and found a desire for explicit coursework on how to teach reading to young learners. Among the seven modules, vocabulary seems to have been receiving the most share of research. Aukrust (2007) examined the effect of the amount, diversity and discoursal complexity of teacher talk on the vocabulary acquisition in young learners, and found significance in all three areas. The Chujo and Nishigaki's (2004) TTC workshops were focused on methods of teaching vocabulary to children. Szpotowicz (2009) examined the main factors that were assumed to influence young learners' vocabulary acquisition, and emphasized that lexical knowledge must be our first priority in TEYL. Unsworth, Persson, Prins, and De Bot (2015) conducted a similar study and elaborated on the lexical aspects that have a greater role in early vocabulary acquisition among young learners. Chujo et al. (2011) compiled a daily life corpus to be used in TEYL. Their work reflects the idea that vocabulary presented to young learners of English must be rooted in the natural language used by the usual companions not extracted from books that do not match the reality of their life. Hirata (2016) followed the same approach and gave the priority of corpus-based vocabulary lessons in TEYL. Graham, Courtney, Marinis and Tonkyn (2017) found that teacher level of training in language instruction had a significant effect on both grammatical and lexical performance of young learners. To see the reaction of our student teachers toward the issue of modulation, one has to see the results section further below.

3. METHOD

Participants

The participants of the study were twenty-two volunteer student teachers who were recruited to be trained in our online teacher training course. They were all BA students of English majors from Kashan and Tehran who felt the need to be instructed on the issue of teaching English to young learners. They had limited experience in teaching and a basic knowledge of TEFL but had not received any academic education with regard to TEYL. The participants were informed that their performance during the course would be later analyzed by the researchers and were also ensured that any information with regard to their claimed competence or observed performance would remain totally confidential. Seven participants were not accessible for the self-assessment phase; therefore, we had to include only the data obtained from the other fifteen.

Design and Procedure

Our study was designed in three phases: 1) Teacher education; 2) Subjective self-assessment; and 3) Objective evaluation. The first phase was accomplished via a 5-session online course. Each session was allocated to one of the five dimensions of TEYL explained above. The teacher educator introduced the issue of the session and engaged in online discussions that involved the entire class. These weekly online sessions were accompanied with offline group discussion during the week. What was called the participants' reaction to the course refers to the nature of their participation in the offline discussions during these five weeks. The oral participation was transcribed and added to the written messages already available in the group. The major themes emerged during the discussions reflected the young teachers' reaction to the presented material. Educational discourse analysis was used to extract the themes. The details are later discussed in the result section.

The second phase sought to tap into the student teachers' perspective with regard to the impact of the course on their competence. The self-assessment questionnaire included 25 items in five sections namely 1) motivation; 2) mediation; 3) maturation; 4) modulation; and 5) manipulation. During the third phase a material development task was given to the participants and their performance was scored by the educator based on the very criteria introduced in the online course. In sum, there were three sets of data: a) original frequency data from the questionnaire along with

their converted interval counterparts; b) interval scores produced by two parallel raters; and c) frequency and textual data recorded and analyzed with educational discourse analysis techniques.

Instruments

We had a self-developed questionnaire to survey the student teachers' conceptions and a rating scale to score their performance on the final task. The reliability of the questionnaire was checked by the use of Alpha Cronbach formula, and the inter-rater reliability of the task scores was calculated through a correlation analysis. The former was based on internal consistency of the questionnaire, and the latter was based on inter-personal consistency of the raters who were using the same scoring criteria. The Alpha Cronbach coefficient was 0.82 which is a within the range of fairly acceptable, and the correlation coefficient was found to be 0.83 which is also acceptable. The numbers show that both instruments were reliable enough for the purposes of this study. Both instruments were used to collect the data needed for the inferential statistics presented in the following section.

4. RESULTS AND DISCUSSION

The results are presented in three sections: 1) The content analysis findings that reflect the student teachers' individual and collective reaction to the course; 2) The chi-square coefficients that show the significantly effective aspects of our TTC as assumed by the student teachers; and 3) The linear regression coefficients that show the extent to which the student teacher's subjectively assumed competence could predict their objectively scored performance.

Textual Data: Reactions to the Course

As explained in the procedure of the study, the oral and written messages exchanged via the WhatsApp group were recorded and color coded for later analysis. The voices were transcribed and added to the written materials that comprised the majority of the textual data. Two lines of investigation were followed in our content analysis: 1) Theme analysis; and 2) Frequency statistics. The former had a qualitative, collective and exploratory nature and was adopted to detect the lines of discussion which were not explicitly mentioned in our defined dimensions, whereas the latter was based on clear counting of specific cases in which a participant was individually discussing one of the five dimensions introduced during the course. It should be mentioned that since the themes are covering issues missed by the course content and the questionnaire, all three of them are directed toward problematic mentalities that function against our defined five dimensions. Table 1 shows the collective results of the theme analysis, and Table 2 shows the individual results of the frequency analysis.

Table 1: Emergent themes in student teachers' collective reaction to problematic issues

	Theme	Semantic Concepts	Pragmatic Context
1	Cultural Dilemmas	Religion, order & discipline, Honesty	Resolving the conflicts
2	Financial Considerations	Salary, Cost effectiveness	Consulting the educator
3	Technical Phobias	Multimedia, Online learning, devices	Complaining about the software

Each theme deals with a certain problem which has been frequently mentioned by different individuals in the course of offline discussions. Every theme is reflected in a series of related key concepts that surface in dialogues and represent important issues to the students. By adding the context column in Table 1, we aimed to avoid a usual negligence in theme analyses which can be called an artificial decontextualization of textual features. It can be argued that the discussed key

concepts and the themes emerging from them need specific contextual clues that build a pragmatic relevance for them. In the absence of informative contextual descriptions, a divergent series of interpretations is possible, and this jeopardizes the scientific unity of the results. For example, when we turn to the theme "Technical Phobias" and look into its key concepts, we have to know if it was brought up as a complaint or a suggestion. The pragmatic context reveals the place of the interlocutors and increases the accuracy of the interpretations. If the pragmatic context in not given to the reader, the textual features can be easily put in an assumed mental context that is far from the reality of the discussions. As mentioned earlier, the themes covered the problematic areas and were recognized collectively. Now we turn to quantitative individual data that represents the share of each participant and the portion allocated to each dimension during the group discussion (Table No. 2).

Table 2: Student teachers' individual shares in discussion over five dimensions of TEYL

| Student NO.	Motivation	Mediation	Maturation	Manipulation	Modulation	Total Contribution
1	124	234	113	342	112	925
2	201	137	200	98	70	706
3	160	198	176	202	245	981
4	164	234	153	124	351	1026
5	321	115	78	200	172	886
6	143	234	132	231	118	858
7	234	125	152	132	362	1005
8	65	166	89	304	253	877
9	254	143	127	321	271	1116
10	124	153	176	156	341	950
11	243	133	145	109	236	866
12	254	123	163	135	152	827
13	199	186	165	263	221	1034
14	234	215	216	117	78	860
15	243	132	221	51	321	968
Total	2638	2157	1993	2345	3121	12254
%	22	18	16	19	25	
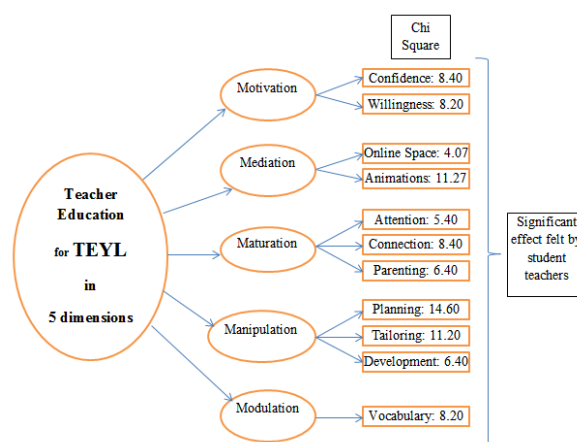
Diagram 1: Student teachers' assumed competence and their significant chi square coefficients

Chi Square: The Assumed Competence

We ran 25 independent non-parametric chi square tests on the frequency data obtained from our 25-item questionnaire. The items were divided into the five dimensions of TTC. Since 14 tests produced insignificant coefficients, here we only focus on the 11 items that did show a meaningful change in the perspective of student teachers. The chi square coefficients in diagram No. 1 show the magnitude of that change as felt and reported by the participants of this study.

The bigger the chi square coefficient, the bigger the effect of TTC on the teachers' competence as felt and reported by them. The participants' answers to the items allocated to the first dimension i.e. motivation shows that from their perspective the contents and discussions of our online course had a significant effect on two aspects of their motivation to teach English to children namely *confidence* to teach and *willingness* to embark on a career with a chi square coefficient of **8.40** and **8.20,** respectively.

With regard to the second dimension i.e. mediation, the two aspects that stand out as statistically significant when one taps into the mentality of the student teachers are *online space* and *animations* with chi square coefficients of **4.07** and **11.27,** respectively. It means that according to the collective opinion of the student teachers, the contents of the online course along with the consecutive offline discussions helped them promote their competence regarding mediation in online TEYL.

The meaningful effect of our TTC on the third dimension, maturation as reported by the student teachers can be seen in the chi square coefficients of three aspects namely *attention* (**5.40**), *connection* (**8.40**) and *parenting* (**6.40**). It means that they think the course has helped them to be more attentive to children, communicate with them and also with their parents. In other words, the student teachers, in a statistically significant way, feel that they have become more mature in their mentality toward young learners of English and their parents.

The fourth dimension of TEYL called manipulation in our study is reflected in the significant results found by the chi square test for the items focusing on three aspects namely *planning* (**14.60**), *tailoring* (**11.20**) and *development* (**6.40**). These numbers tell us that the contents of our online sessions accompanied by the following offline discussions were successful in creating a self-improvement in the student teachers when it comes to issues traditionally discussed under titles such as curriculum planning, syllabus design and material development.

The fifth dimension of our TTC for TEYL focused on modulation of teaching activities across language skills and components. The results of the chi square test on the relevant items show that

it is only *vocabulary* that stands out significantly as reflected in a coefficient of **8.20**. In other words, the student teachers did not find the contents of the course significantly effective in promoting their competence with regard to the other two components or all the four language skills. It was only the discussions revolving around the issue of teaching vocabulary to young learners of English that had a meaningful and positive effect on the relevant competence of the student teachers as reported by them via the self-assessment questionnaire.

Regression Analysis: The Predicted Performance

Having elaborated on the non-parametric results, now we turn to the parametric findings. As explained in the method section, each participant received a score based on their performance on the final task. The scoring criteria matched those introduced during the course. These scores provided us with an objective base against which one could compare the subjective statements of the student teachers. The multimedia material produced by the trainees showed to what extent their claims with regard to the efficiency of the course are reliable. In others words, we wanted to know if the claims of efficiency (reflected in the questionnaire) can predict the real outcome (presented in the task) in a statistically significant manner. To this end we used a regression analysis, which is based on correlation coefficients. To be able to calculate correlation coefficients, we first had to convert the Likert scale data into interval numbers so that this new score could represent the students' perspective with regard to the effectiveness of the course. The stronger the relationship between the assumed competence and the observed performance, the more reliable our results would be. Table 3 below presents the results of the regression analysis.

Table 5 has two main sections: the preliminary ANOVA and the main test that is a Linear Regression. The former has to be significant; otherwise, the values of the main test are not applicable. As it can be seen the ANOVA did produce a significant F value, and this allowed us to proceed with the regression results which were also significant in their own terms. To put the latter into perspective, one can imagine a line equation with the assumed competence as the independent variable, and the observed performance as the dependent variable, or in mathematical terms, x and y, respectively. The significance of t=3.69 implies that the variation in students' performance on the task could be predicted to an acceptable extent (58%) by relying on the impact they had attributed to the course. In more clear terms, what the student had imagined to have learned from our course had been translated into their actual performance by 58 percent. Other hidden factors and erroneous variables would naturally account for the unpredicted portion of their performance.

Table 3: Prediction of Student teachers' performance by their assumed competence

ANOVA		Regression		
F	Sig	Beta	t	Sig
18.04	0.00	0.58	3.69	0.00

5. DISCUSSION

The main results can be summed up and compared against the available relevant literature below.

According to the student teachers' collective reaction percentages in our results, the maturation dimension of TEYL was the least discussed. This is consistent with the findings of a series of studies in the literature (see Cameron, 2003; Cao, 2019) that report the underestimation of communicative and affective aspects of TEYL in teacher training programs.

To look outside the box, we tried not to be confined to our own defined dimensions hence tracing student teachers' emergent themes in textual discussions to find the underrepresented and neglected areas that contribute to their teaching anxiety. We realized that young teachers who embark on a TEYL career, in spite of knowing the right path, might remain pinned down by their worries and fears in three main areas: a) cultural dilemmas, b) financial consideration and c) technical phobias all of which are considered main issues in the literature. Pejović (2013) suggest investing in intercultural awareness to resolve the cultural conflict in TEYL. Sowa (2017) argues that our demands from the young teachers should match their financial capacity; of course, it would be effective to promote the latter so that a higher quality of teaching can be expected. Lam (2000) recognized technophobia as a problem worth tending to, and explains how one can help young teachers become technophiles in time. We suggest that the interaction of these problematic areas and their contribution to teaching anxiety be investigated further in future studies.

The most central focus of our study was the relationship between the student teachers' assumed competence and actual performance. Throughout the paper, we kept emphasizing that it is "assumed" competence not the competence itself. In accordance with studies such as Wissink and Starks (2019), Strakova (2015), and Sad (2015), we delimited ourselves to the teachers' perception of their teaching competency because we were interested to tap into the teachers' subjective self-image rather than their repertoire of knowledge. The results showed that what student teachers think they can do in TEYL significantly predicts what they can actually do for the young learners.

6. CONCLUSION

In Iran, TEYL needs teacher education programs more than ever. Although the need is felt globally, meeting it requires local efforts. Therefore, being already involved in this area for several years and with an action research agenda in mind, we held a TTC to educate young student teachers and analyzed the output via descriptive and inferential statistical methods.

We were interested to see what collective and individual reactions we would receive from the student teachers. We also wanted to see if what they think and claim with regard to the effect of our course on their teaching competence can predict their actual performance at least to an acceptable degree. To these ends, we collected textual and numerical data from 1) discussions during the course; 2) answers to our questionnaire; and 3) teachers' performance on a final task. The idea was to first educate the student teachers, then tap into their mentality and finally observe them in action. The results showed that the teachers' performance is meaningfully dependent on their mentality with regard to the priorities of teaching English to young learners and also on their self-image as formulated along the five dimensions of TEYL. The findings can effectively contribute to the current practice of English teacher education in Iran. The future studies can focus on any of the dimensions and delve into their internal mechanism. In addition, transferring this experience into a wide range of educational contexts can yield more reliable results with a higher degree of generalizability and validity.

References

Ahmad, N., & Samad, A. (2018). Metaphors as proxies for identity: A case study of a teaching English to young learners (TEYL) teacher. *The Southeast Asian Journal of English Language Studies, 24*(4), 143-157. https://doi.org/10.17576/3L-2018-2404-11

Aukrust, V. (2007). Young children acquiring second language vocabulary in preschool group-time: Does amount, diversity, and discourse complexity of teacher talk matter? *Journal of Research in Childhood Education, 22*, 17-36. https://doi.org/10.1080/02568540709594610

Cameron, L. (2003). Challenges for ELT from the expansion in teaching children. *ELT Journal, 57*(2), 105-112. https://doi.org/10.1093/elt/57.2.105

Camlibel-Acar, Z. (2017). Pre-service teachers' perceived readiness and willingness to teach English to young learners. *Studia Anglica, 226*(7), 21-30.

Cao, P. (2019). Teaching English to young learners in Vietnam: From policy to implementation. *The Asian Journal of Applied Linguistics, 6*(1), 96-104.

Chujo, K., & Nishigaki, C. (2004). Creating E-learning material to teach essential vocabulary for young EFL learners. *IWLeL: An Interactive Workshop on Language e-learning*, 35-44.

Chujo, K., Oghigain, K., Uchiyama, M., & Nishigaki, C. (2011). Creating corpus-based daily life vocabulary for TEYL. *Asian EFL Journal, 54*, 30-58.

Copland, F., Garton, S., & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly, 48*, 738-762. https://doi.org/10.1002/tesq.148

Dorsah, P. (2021). Pre-service teachers' readiness for emergency remote learning in the wake of COVID-19. *European Journal of STEM Education, 6*(1), 1-12. https://doi.org/10.20897/ejsteme/9575

Graham, S., Courtney, L., Marinis, T., & Tonkyn, A. (2017). Early Language Learning: The impact of teaching and teacher factors. *Language Learning, 67*(4), 922-958. https://doi.org/10.1111/lang.12251

Hirata, E. (2016). The development of multi-modal corpus tool for teaching English to young learners. *Fukuoka Jo Gakuin University Bulletin, 2*, 19-32.

Hussein, R. (2014). A profile of exemplary teachers of English for young learners at the elementary school. *Jurnal Pendidikan Humaniora, 2*(4), 311-321.

Lam, Y. (2000). Technophilia vs. technophobia: Why second-language teachers do or do not use technology in their classrooms. *Canadian Modern Language Review, 56*(3), 389-420. https://doi.org/10.3138/cmlr.56.3.389

Nunan, D. (1997). Designing and adapting materials to encourage learner autonomy. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 192-203). London: Longman.

Pejović, I. (2013). Cultural pluralism, multiculturalism and intercultural awareness in TEYL: The eyes of the same color. In V. Savić & I. Ćirković-Miladinović (Eds.), *Integrating culture and language teaching in TEYL, conference proceedings No. 16* (pp. 19-26). Jagodina: City Press.

Pim, C. (2013). Emerging technologies, emerging minds: Digital innovations within the primary sector. In G. Motteram (Ed.), *Innovations in learning technologies for English language teaching* (pp. 15-42). London: The British Council.

Ratminingsih, N., Mahadewi, L., & Divayana, D. (2018). ICT-Based Interactive Game in TEYL: Teachers' Perception, Students' Motivation, and Achievement. *International Journal of English Teaching, 13*(9), 190-203.

Rich, S. (2014). Taking stock: Where are we now with TEYL? In S. Rich (Ed.), *International perspectives on teaching English to young learners* (pp. 1-22). UK: Palgrave.

Sad, S. (2015). Do prospective classroom teachers perceive themselves as effective and willing to teach young learners of English? *Educational Sciences: Theory & Practice, 15*(5), 1257-1269.

Strakova, Z. (2015). The perception of readiness for teaching profession: A case of pre-service trainees. *Journal of Language and Cultural Education, 3*(1), 32-42.

Szpotowicz, M. (2009). Factors influencing young learners' vocabulary acquisition. In M. Nikolov (Ed.), *Early learning of modern foreign languages: Processes and outcomes* (pp. 195-211). Clevedon, UK: Multilingual Matters.

Unsworth, S., Persson, L., Prins, T., & De Bot, K. (2015). An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics, 36*, 527-548. https://doi.org/10.1093/applin/amt052

Wissink, B., & Starks, S. (2019). Elementary teachers' perceptions of preparedness to teach English language learners. *Educational Research and Reviews, 14*(10), 349-357.

Woolfolk, A. (2008). *Educational psychology: Active learning edition*. Boston, MA: Pearson.

The Effect of Using a Mobile Game-Based Application on Iranian EFL Learners' Pronunciation: Exploring their Motivational Perception

Fatemeh Shirmardi[1], Ali Roohani[1,*] iD, and Bashir Jam[1]

[1] Shahrekord University, Shahrekord, Iran

Abstract

Mobile applications have emerged as promising tools contributing to second/foreign language (L2/FL) learning. This study examined the effect of playing Spaceteam ESL, a game-based mobile application, on FL (English) pronunciation, which is less explored in L2 education. It also explored motivational perceptions of FL learners regarding the game as an effective pedagogical tool. To these ends, 40 female English as foreign language (EFL) students at low-intermediate level from 2 high school classes in Iran participated in the study after taking a placement test. They were assigned into 2 groups: experimental and control groups. In order to measure their pronunciation ability regarding some target English sounds, a researcher-made pronunciation test was administered to the participants as a pretest. The experimental group used the game, in addition to other activities, in the class for 10 weeks; the control group was involved with non-gaming activities. At the end of the treatment, the pronunciation test was given as a posttest. To explore the EFL students' perception of the game-playing, the students in the experimental group were interviewed. The result of t-tests revealed that the experimental group obtained better outcomes than the control group in the pronunciation. The analysis of the semi-structured interview data revealed the students' positive attitude in playing the Spaceteam ESL game and their high motivation for using such games. The findings provide implications and suggestions for students, teachers, and course designers in L2 pedagogy.

1. INTRODUCTION

As evidenced by both anecdotal evidence and a number of studies (Pallier, Colomé, & Sebastián-Gallés, 2019), correctly perceiving and producing the sounds of a second/foreign language (L2) is difficult. As Hismanoglu (2006) states, L2 teachers should give great importance to teaching pronunciation. In L2 teaching and learning, pronunciation plays a crucial role for communication, and it requires the sound system comprehension and production of target sounds (Haghighi & Rahimi, 2017). However, in L2 research, the instruction of pronunciation has taken

---

*** Corresponding Author:** Ali Roohani, Department of English, Faculty of Letters and Humanities, Shahrekord University, Shahrekord, Iran. **Email**: roohani.ali@gmail.com

a back seat, compared with such aspects as syntax and morphology (Deng et al., 2009). This secondary role has directed some researchers (e.g., Derving & Munro, 2005) to claim that teaching pronunciation is commonly neglected. According to Hedge (2000), L2 researchers either utterly omitted pronunciation from major publications, or gave limited attention to the topic. However, instruction on pronunciation, as Lee et al. (2019) put it, can be as helpful as grammar, vocabulary, and pragmatics. Moreover, recent research in L2 education has focused on how technology can be integrated most effectively in L2 courses, including pronunciation courses (Fouz-González, 2018). Researchers are interested in the use of recent technology as a promising tool, because it can improve presentation styles and make instructional materials more psychologically and physically available (Pennington, 1996), provide students with individualized practice and with many tries, and offer them automatic and immediate input on their actual performance (Fouz-González, 2018).

According to Hsu (2016), the ever-going technology development over the past decades has changed language learning and teaching landscape. Greatly developed digital literacy skills of learners and great range of technological applications permit a new way to teaching foreign languages in general, and teaching L2 pronunciation in particular. In integrating digital technologies into learning/teaching environments, some Computer-Assisted Language Learning (CALL) and Mobile-Assisted Language Learning (MALL) researchers develop or use game-like practice or activities for learning of language. Students involving in such practice are given activities and expected to achieve them by working in CALL and MALL environments such as virtual worlds (Zheng, Young, Brewer, & Wagner, 2009) or by playing a mobile game (Reinders & Wattana, 2014). Playing games can help students gain knowledge of the world (Coyne, 2003).

Investigations on mobile game-based learning can also focus on the motivational effects of the methods. In L2 learning, motivation is crucial because it provokes, directs, and keeps better performance in learning (Boo, Dörnyei, & Ryan, 2015; Chang, 2005). Motivated students devote more effort to their L2 learning; they persevere more even if the long process of L2 learning is difficult and allocate more time and energy to doing so (Ushioda, 2013). Some MALL scholars (e.g., Ena, 2003) have asserted that mobile applications can enhance students' motivation for L2 learning. As an instance, Sandberg, Maris, and de Geus (2011) claim that learners are motivated to utilize a mobile phone application to learn English outside the class even if they are not required to do so. More specifically, using various interesting activities through games may improve the learners' English pronunciation ability (Nurhayati, 2015). Games might also be effective because they lower learners' stress and give them the opportunity for real communication. Despite their benefits, the integration of digital games into formal L2 education is still rare and different problems have been recognized. The most notable problem is that there is little acceptance of games as pedagogical means among most of L2 teachers (Wastiau, Kearney, & Van den, 2009). Some teachers consider games as leisure time activities without any instructional value.

Facing the demanding work environment due to the needs of English learning in the educational context of Iran, EFL teachers might struggle to find alternative methods to foster development in learners' pronunciation and motivate learners in an engaging and playful atmosphere. In this way, game-based learning may be promising. Given the potential effect of game-based instruction on L2 pronunciation and motivation, the present study sought to investigate the possible effect of a game, called the *Spaceteam ESL* on L2 learners' pronunciation and explored their motivational perceptions of the game.

2. LITERATURE REVIEW

Teaching Pronunciation

Historically, pronunciation as an aspect of L2 learning developed with the audio-lingual method and situational language teaching, both of which considered it as a vital language component (Morley, 1991). These two approaches used articulatory explanations, imitation and pattern drills, with a major emphasis on correction. As Morley (1991) asserts, they abandoned the method of articulatory explanations and replaced it by a more communicative and functional method. In the 1960s, the importance of teaching pronunciation was gradually questioned. At this time, many curricula put less emphasis on pronunciation or drop it from syllabi (Morley, 1991). An example is the cognitive approach that put great emphasis on grammar and vocabulary rather than on pronunciation (Celce-Murcia, Brinton, & Goodwin, 1996). Later, in the 1970s, teaching pronunciation was discussed theoretically and extensively. The 1980s was characterized by a new interest in pronunciation teaching, influenced by the notions explored theoretically in the previous decade. One of the methods that gave importance of pronunciation was the communicative approach. Giving a great emphasis to effective communication, the communicative approach focused on the need to help L2 students attain a certain level of pronunciation skills, that is, a threshold above which the communication is not distorted by misunderstandings caused by an incorrect pronunciation (Celce-Murcia et al., 1996).

Mobile-Assisted Language Learning

In the 1950s, the advent of audiolingual method brought about the tremendous utilization of language laboratories in academic environments (Salaberry, 2001). Along with the development in technology, as Chinnery (2006) maintains, the laboratory was increasingly substituted with drill-based computer-assisted instruction in the 1960s, which was overshadowed, decades later, by multimedia, interactive, and intelligent CALL. In the 1990s, according to Chinnery (2006), the widespread endorsement of the Internet gradually promoted the expansion of Computer-Mediated Communications (CMCs), and technologies with larger capabilities (e.g., emails, videoconference, and network communication) started to emerge. These rising technologies have been characterized by their ease-of-access, feasibility, mobility, and transportability.

During the chronicle of educational technology, there have been tendencies dealing with capabilities and size of technologies, which support instruction in the class. Small technological devices such as mobiles to able to do as much as, and even more than, large computers. Characteristics such as mobility, portability, and availability have made the expansion of mobile language learning applications and tools possible (Al-Zahrani, 2015), resulting in the development of MALL. Some researchers (e.g., Viberg & Grönlund, 2012) categorize MALL as a sub-branch of m-Learning (mobile-learning), while others believe that it is a subset of the CALL domain. According to Joseph and Uther (2006), what makes MALL extend the CALL field into everyday activities is the mobility feature of MALL. As Viberg and Grönlund, (2012) contend, MALL is a sub-branch of m-learning that has progressively attracted the interest of some researchers. For instance, Huizenga, Admiraal, Akkerman, and Dam (2009) and Su and Chengt (2015) have developed game-based learning activities in education and gradually the term gamification has emerged in the literature. In fact, gamification is a new term referring to the use of game mechanics and game design elements in non-game contexts with the aim of engaging individuals and solving problems. It tends to be a solution for involving them in sustainable behaviors, like exercise, practice, and instruction (Girard, Ecalle, & Magnant, 2013). Based on research, gamifying a course can be very helpful for school students (Sandberg, Maris, & deGeus, 2011). By keeping the

motivational power of games into consideration and implementing it to the motivational problems in teaching, more learning achievement can occur (Prensky, 2001).

Many game applications have been designed in recent years, including *Conjugation Nation, Edugame, Frequency 1550, and Massively Multiplayer Online Role-Playing Game* (*MMORPG*). The *Frequency 1550,* as an example, is a Dutch game working in cultural and social domain. It is a mobile game in which pupils acquire historical knowledge about medieval Amesterdam. Groups of four or five are formed to obtain citizenship in the Amesterdam city via getting the required points or days of citizenship (Huizenga, Admiraal, Akkerman, & Dam, 2009). *MMORPG* is a type of popular digital game played by many players in a complicated environment which needs a substantial amount of player interaction to progress in the game (Reinders and Wattana, 2015). Another popular game application is *Spaceteam ESL*. As Grimshaw and Cardoso (2018) explain, *Spaceteam ESL* is a portable and easily accessible mobile application with such features as CMC, practice of common terms, time-dependent interaction. This game can create a learning environment for learners to practice English. In their study, Grimshaw and Cardoso (2018) investigated the possible effect of the *Spaceteam ESL* on oral fluency and perceptions of 20 Canadian English as Second Language (ESL) learners. They observed changes in the Canadian ESL students' oral fluency (i.e., judges' ratings and syllables produced each minute). Moreover, the gameplay positively influenced the ESL learners' willingness to communicate and reduced their anxiety.

Also, there has been an increasing demand to use technology with special emphasis on pronunciation. For instance, Fouz-González (2020) conducted a research to explore the potential of a digital application, *English File Pronunciation (EFP),* to help 47 advanced Spanish EFL learners to improve their production and perception of some sounds that were likely to be fossilized in the interlanguage of the Spanish EFL learners. The results of pre- and posttest revealed that although the difference between the group that used the mobile application and the one that did not use it was not statistically significant for the target sounds, the application improved the learners' production and perception of the segmental sounds.

There are a number of studies on MALL and L2 skills, such as listening (e.g., Liu & Chu, 2010; Miangah & Nezarat, 2012) and speaking (e.g., Grimshaw & Cardoso, 2018), L2 interaction (Reinders and Wattana, 2015), and academic achievements (e.g., Su & Chengt, 2015; Partovi & Razavi, 2019). Despite tremendous empirical studies done in the MALL domain, the effect of game-based applications on L2 pronunciation domain has given less attention. A close look into literature indicates that there is little empirical research on the effect of *Spceteam ESL* application on English pronunciation, which is the focus of the present study.

Motivation

Deriving from the Latin word *movere*, motivation is defined as what moves an individual to make a certain decision, to engage in an action, to spend effort, and to persevere in action (Dörnyei & Ushioda, 2011). Research on language learning motivation has witnessed more than 50 years of ongoing history and has observed different phases of progress, with major milestones such as Gardner and Lambert's (1972) views of motivation. Gardner and Lambert highlighted the significance of attitude and culture in language learning and introduced integrativeness, a notion which relates to the desire to integrate and learn more about another culture, and instrumentality, which relates to the functional value of learning a language (i.e., job purposes). Gardner's view and his social approach were so dominant than other motivation views existing before the 1990s. However, Gardner's theory failed to take the new conception of social identity into account. Thus, new studies on motivation started to emerge. Cognitive traits proved to have great importance in

educational settings, and L2 researchers employed them in L2 teaching and learning. This way, the cognitive viewpoints, namely, attribution and the self-determination theories, emerged.

These emerging theories and approaches stressed the notion of intrinsic and extrinsic motivation (Ghapanchi, Khajavy, & Asadpour, 2011). Intrinsic motivation was described as the fulfillment of learning for personal desire and satisfaction, wheras extrinsic motivation was associated with external regulations like rewards, choiceful behavior, and forced rules (Dörnyei, 2001). The reconceptualization of the L2 motivation has occurred in recent years. Now, the recent views of motivation embrace both the learner's social context and view of the self as components linked to L2 motivation. In actual fact, because there was dissatisfaction with the more traditional model of L2 motivation, the L2 motivation theories have been reconceptualized. As Ebrahimzadeh and Alavi (2016) maintain, the traditional models considered motivation as a linear phenomenon while motivation results from a sequence of multiplex interactions. Moreover, they took a reductionist perspective toward motivation, describing a series of specific variables that would predict motivation. The socio-dynamic views, however, considered the complication of the process of the L2 motivation and its circular change in dynamic interaction with a variety of contextual, internal, and social factors whilst taking into account the greater complication of language learning and language use in the globalized world. That is to say, motivation could not be defined and identified without recognizing the interactions and relationships between the various fundamental motives in the context. Moreover, persons' thoughts and their perceptions toward the context change over time. This issue emphasizes that motivation is a dynamic phenomenon.

Taking the importance of motivation into account, L2 researchers are continuously involved in implementing effective and interesting teaching strategies and methods to encourage language learners to learn and enhance their motivation. Naturally, some scholars (e.g., Sudarmilah et al, 2020; Cho & Castaneda, 2019) have made attempts to incorporate the mobile devices in language teaching practice. Cho and Castaneda (2019), for instance, investigated if any changes would occur in Spanish learners' affective and motivational engagement in L2 after engaging in game-like activities with the *Conjugation Nation* application. *Conjugation Nation* is a mobile application developed to promote grammar learning in the Spanish language. To that end, 82 learners of Spanish in six classes in a university in the US took part in the study during four semesters. The results indicated that the game-like practice with the *Conjugation Nation* was partially useful in increasing positive affect. The participants largely noted that the activities were helpful and enhanced their enjoyment of learning Spanish. Their study supports the contribution of gameplay in language learning. Also, Kétyi (2013) asserts that learners can gain additional learning time out of the classroom, be motivated, and develop their language learning if their teachers incorporate the mobile devices in language teaching practice. Further research shows that integrating mobile game-based applications in learning environment has the potential to enhance learners' motivation and encourage language learning. For instance, Ebrahimzadeh and Alavi (2016) carried out an experimental study to examine the effect of utilizing digital games for learning English vocabulary as a means to motivate EFL learners and create enjoyment. They found that the Iranian EFL learners who were engaged in the digital gameplay enjoyed learning vocabulary, which considerably improved their motivation to learn English. Likewise, in a prior study, Schwabe and Göth (2005) conclude that technology could enable learners to engage into a mixed reality and, consequently, provide motivating learning experiences. According to them, mobile games can push learners into a state in which they are mentally prepared for learning.

Despite the enormous research done in MALL, there is scarce evidence of pedagogical implementation of the game-based application in L2 pronunciation development and motivational outcomes, particularly in the EFL context of Iran. Many Iranian high school EFL students,

particularly at the low level of proficiency, have problems with the pronunciation of some English sounds, such as [w], [θ], [ð], [uː], [ʊ], [iː], and [ɪ], which do not exist in Persian. It is then promising to look into the effect of using game-based instructions on high school students' English pronunciation accuracy and motivation, and fill the existing gap in the L2 gamification and pronunciation studies. The current study was intended to examine the effect of playing *Spaceteam ESL,* a mobile game-based application on low-intermediate high school students' English pronunciation accuracy. Also, it sought to explore the high school students' views about the *Spaceteam ESL* game application as a means to enhance their motivation. In line with these objectives, this study provided answers to the following two questions:

Does playing the *Spaceteam ESL* game improve English pronunciation accuracy of Iranian low-intermediate high school EFL students?

What are Iranian low-intermediate high school EFL students' perceptions of the *Spaceteam ESL* application as a means to enhance their motivation?

3. METHOD

Participants

The study was carried out with a sample of 40 Iranian EFL students from 2 classes in Andika, Khouzestan, Iran. They were female high school EFL students aged between 17-19 years old, with a mean age of 18. They were studying English in the high school and had not taken any course on learning pronunciation in language institutes. They were selected after taking a language placement test. They were at the low-intermediate level. Half of the participants were in the control group (n = 20) and the other half were in the experimental group (n = 20).

Instruments

The instruments were a placement test, a pronunciation test, and a semi-structured interview. In order to ensure the homogeneous entry of the students, a placement test, Quick Placement Test (QPT, 2001), was initially administered to measure the participants' English language level. The test consisted of 60 items in multiple-choice format, measuring vocabulary and grammar knowledge.

A researcher-made pronunciation test was developed to measure students' pronunciation ability before and after treatment. It was used as the pretest and posttest. This test focused on the segmental features such as sounds, [w], [θ], [ð], [uː], [ʊ], [iː], and [ɪ]. which were absent in Persian. The pronunciation test had 45 items and consisted of three parts: In part one, the participants listened to an audio file and marked the word they heard in the audio. In part two, the participants listened to several sentences and selected the correct pronunciation of the underlined words. In part three, the participants were required to read some words (minimal pairs such as 'slip' vs 'sleep' or 'fill' vs 'feel') with correct pronunciation with their voices being recorded. The total score for the pronunciation test was 45, and time allocation was 30 minutes. The content validity was checked by experts' judgments. The test content corresponded to pronunciation features of their high school textbook, *English for Schools, Vision 3* (Alavi Moghaddam, Kheirabadi, Rahimi, & Davari, 2020) taught in high schools in Iran. The test received coverage of the segmental features of pronunciation covered in the program. As to the reliability of the test, the test-retest reliability of the test in a pilot study with 30 high school students, who were similar to the EFL participants, was high (.75). Also, the Kuder-Richardson 21 index for parts 1 and 2 of the test was found to be acceptable (0.80). In addition, the interrater reliability for the part 3 of the test was high (.98).

Figure 1: A screenshot of the two phases in the *Spaceteam ESL* gaming application.

In order to explore the participants' views about the *Spaceteam ESL* application as a pedagogical means, a semi-structured type of interview was conducted with the participants in the experimental group. It included several open-ended and yes/no questions such as "How do you evaluate your experience of using the *Spaceteam ESL* game?", "How did you find the method employed in this course? Was it interesting or not? Why?", and "Did the class encourage you to keep going in your improving pronunciation? If so, how?". Two experienced instructors to have feedback on their appropriateness, clarity, and relevance reviewed the questions.

Procedure

A pretest-posttest control group design was employed; it consisted of two groups: treatment and control. At the beginning, the placement test was used in order to check the English language level of the participants (n = 40) and ensure their homogeneity at the entry level before they receive the instructions of the study. Their scores fell within 30-39 (low-intermediate level). The participants in one class were assigned as the experimental group (n = 20) and the participants in another class were assigned as the control group (n = 20). The experimental group were exposed to the game (*Spaceteam*) once a week for ten weeks. At the start of each session, they were engaged in the gameplay for 30 minutes. *Spaceteam ESL* is a double-mode game. In the first mode, the students/players were provided with a list of vocabularies consisting of the most-frequently-used vocabularies on their mobile screen. An automated word producer produced the words in native English. The students were then expected to repeat the same vocabularies. In the second phase, the same lexicon in the word list appeared in the game while the students were playing. Figure 1 displays the two phases of the game in the application.

In order to make the participants understand the course objectives and get familiar with the *Spaceteam ESL*, they had a training session. By using this application, the experimental group was set up for students to play in teams of four. Then, they installed the *Spaceteam ESL* application on their smartphones or tablets in order to keep a spaceship running. The participants in the experimental group interacted with each other as a group to accomplish the game. As the game level increased, the word complexity, in terms of frequency level of words and difficulty level in pronunciation increased. They interacted to send and receive instructions in a limited amount of time. If there were any misunderstanding, the student-interlocutor regularly repeated the wrong word or mispronounced vocabulary to ensure if it was the target one. The written instructions on the screen provided more clues for interpreting their classmates' response. The student-

interlocutors were able to give peer input on what was read and heard. The interlocutor was able to identify and pronounce the correct form again.

The students in the control group took part in traditional activities commonly used in such classes, including listening to the target sounds and imitating those sounds. In this way, the teacher provided them with clear distinction between different aspects of pronunciation such as the position of the tongue, manner and place of articulation. Moreover, they completed paper-based supplementary materials in which they practiced listening to the sounds (i.e., vowels and consonants), producing those segments, and completing correct phonetic symbol for vowel and consonant segments.

After 10 weeks, the pronunciation test was given to the EFL students in the experimental and control groups as a posttest to see whether the game-based instruction had a significant effect on L2 pronunciation after the intervention. For the posttest, the test item order was reshuffled. Then, the students in the experimental group took part in the semi-structured interview.

The investigation consists of two distinctive phases: quantitative followed by qualitative. Quantitative data collected through the pronunciation pretest and posttest were analyzed by SPSS (version 25) and t-tests were run to assess the group difference on the pretest and posttest. In order to interpret the qualitative data, the results of the interview were transcribed and coded and the researchers identified themes.

4. RESULTS

The First Research Question: The Effect of the *Spaceteam ESL* Game

The first research question was formulated to gain information about the role that using the *Spaceteam ESL* application played on improving the pronunciation accuracy of Iranian EFL students at low-intermediate level. Table 1 shows the descriptive statistics results regarding pretest and posttest pronunciation scores obtained by the participants in both groups.

As Table 1 shows, the pronunciation mean score of the students in the control group ($M = 10.29$, $SD = 2.61$) was not greatly different from the pronunciation mean score of the students in the experimental group ($M = 11.10$, $SD = 2.45$). Also, the skewness and kurtosis data revealed that the assumption of normality was observed in the distribution of the pronunciation scores of the two groups. However, the score differences between the two groups and the assumption of parametric test was tested statistically. The Kolmogrov-Smirnov tests of normality on the pronunciation pretest and posttest scores showed no significant violation in both groups. The scores were normally distributed for both groups in the pretest, as found to be D (20) = .138, p >.05 in the control group and D (20) = .128, p >.05 in the experimental group. The data of the control and experimental groups in the posttest were also found to be suitable for parametric tests, D (20) = .128, p >.05 and D (20) = .129, p >.05 respectively.

Table 1: Descriptive Statistics of Pretest and Post Pronunciation Scores

Group	Variable	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
Control	Pretest	8	13	10.29	2.61	-.32	.51
	Posttest	13	26	21.72	1.24	-.27	.52
Experimental	Pretest	9	13	11.10	2.45	-.27	.51
	Posttest	19	39	34.86	1.53	-.28	.52

Table 2: Independent t-Test on the Pronunciation Pretest Mean Scores Between the two Groups

| | F | Sig. | t | df | Sig. | Mean Difference | 95% Confidence Interval | |
							Lower	Upper
Equal variances assumed	21.22	.720	.151	38	.881	.81	2.05	2.24
Equal variances not assumed			.151	41.5	.881	.81	2.06	2.25

Table 3: Independent t-Test on the Pronunciation Posttest Mean Scores Between the two Groups

| | F | Sig. | t | df | Sig. | Mean Difference | 95% Confidence Interval | |
							Lower	Upper
Equal variances assumed	1.307	.003	9.21	38	.000	13.14	1.21	1.50
Equal variances not assumed			.151	51.48	.000	13.14	1.21	1.50

Also, independent t-tests were run on the pronunciation mean scores of the control and experimental groups in the pretest and posttest to examine whether the difference between two groups in the pronunciation means were statistically significant. The results of t-tests are illustrated in Tables 2 and 3.

As Table 2 indicates, there was not a statistically significant difference in the pronunciation means of the students between the two groups in the pretest, $t(38) = .151$, $p = .881$, meaning that there was no significant difference between the two groups in the pronunciation before instruction. However, the results demonstrated that that there existed a significant difference between the two groups in the pronunciation achievement after instruction. As Table 3 illustrates, a statistically significant difference was observed in the pronunciation means of the students who used the *Spaceteam ESL* application and the students who did not use a game application, $t(38) = 9.21$, $*p < .05$. The experimental group was better than the control group in the pronunciation gains in the posttest. Also, as the results of the paired t-test shows (see Table 4), the experimental group had a significant gain in pronunciation accuracy from the pretest phase to the posttest phase. They received significantly higher pronunciation scores in the posttest, $t(19) = -20.02$, $*p < .05$.

Table 4: Paired t-Test on the Pronunciation Pretest and Posttest Mean Scores

| | | | | Paired Differences | | | | |
| | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Variance | | t | df | sig. |
				Lower	Upper			
Pretest-Posttest	-23.76	.92	.40	-1.46	-2.38	-20.02	19	.000

The Second Research Question: Perceptions of the *Spaceteam ESL* Game

In order to address research question 2, thematic analysis of the transcribed data from the experimental students were conducted. Six themes emerged inductively from the transcription of interviews. The main themes include atmosphere, social interaction, encouragement, facilitation, anxiety, and challenge.

Atmosphere. The interviewees mentioned the fun atmosphere of the classroom. Some students reported that the attractive nature of the classroom encouraged the learners to pursue and finish the class time without feeling of discomfort, whereas the classrooms (traditional ones) discourage the learners: One student (Maryam, 17 years old) said, "I think the application was great because the students prefer to use a mobile phone. This program is very efficient ... the pressure is high when we have merely the textbook-based teaching. Another student (Fatemeh, 17 years old) explained, "This program was very effective in teaching the language since it created a kind of fun and, I think, all of the students liked learning English by using the application … it is attractive." Also, some participants said that the course was a new experience in English classes. One student (Zahra, 17 years old) mentioned: "It is very fun and interesting, … and I could not feel the passing of the time while playing the game. Last year, the English class was boring, but this class was a new experience … I hope games can be used in the future."

Social interaction. The participants also spoke of collaboration and peer interaction aspect of the *Spaceteam ESL*. Interaction with peers in a context free of anxiety was the reason that the participants referred to. One of the students (Somayeh, 18 years old) stated, "The game helped me to interact with other students and pronounce the sounds native-like". The social interaction aspect of the *Spaceteam ESL* application assisted the participants in the experimental group to improve their pronunciation. As a female student (Maryam, 17 years old) explained: "I liked the game. I like to speak with other people. The game helped me interact with my classmate who was at the same level and improved my pronunciation". This issue was highlighted by Kobra (a 17-year-old student): "The game experience facilitated interaction with peers and created a good learning environment. It helped me have communication with my classmate in the class and, as a result, practice my pronunciation."

Encouragement. Some of the interviewees demonstrated positive attitude towards the game-based instruction and believed that *Spaceteam ESL* encouraged them to learn English better and improve their pronunciation. One student (Parinaz, 18 years old) stated, "I think some students, including me, did not like learning English, but the game-based teaching method has encouraged us to learn English and improve our pronunciation." Other students also viewed the game-based instruction as an encouraging factor: "Such games are interesting and encouraging for students." (Hamideh, 17 years old). Another student specified her reasons as follows: "Playing the game encouraged me a lot to focus on my pronunciation improvement. Such applications can encourage students to learn pronunciation and internalize the teaching content." (Giti, 19 years old)

Facilitation. Some of the students in the experimental group noted that the native-like pronunciation features such as accent observed in the game application could facilitate the process of pronunciation acquisition/learning. One student (Maryam, 17 years old) said, "I love this game-based application … the process of pronunciation acquisition is facilitated … A good way to be active is using such games. It can be used outside the class to improve my accent and vocabulary pronunciation ." Another student viewed oral output as a factor that helped her to facilitate the process of pronunciation acquisition: "It feels good to play such games... we practiced the pronunciation of sounds and new words we did not know ... the oral output provided by the game helped me facilitate the difficult process of learning pronunciation."

Anxiety. Some students also claimed that the game reduced the degree of their anxiety, enabling them to exercise their pronunciation with their classmates without stress. A Low level of students' anxiety, as reported by some participants, was due to teamwork and enjoyable atmosphere in the class during the course. This was explained by one student (Marziyeh, 18 years old): "Most of the time I felt kind of stress and anxiety in English classes before, but this course was different. There was laugh, teamwork, and enjoyment during the course". In the same way, Maryam (17 years old) mentioned, "Sense of laugh and group work made me feel good and reduced my anxiety ... I practiced the correct pronunciation of some sounds."

Challenge. Some of the interviewees (30%) reported that game-based instruction was a challenge for them. They reported that using such applications in the class was challenging and exciting though it was time-consuming. They believed that it involved a lot of effort and time. They said that playing the game was a lengthy process, but it was motivating. This issue was described by one student (Malihe, 19 years old): "Although I was faced with a lack of time in the class when I was engaged with the game, I liked it…it was stimulating and kind of challenge for me ... I wanted to complete the game". Another student (Shadi, 17 years old) commented, "I liked it ... it was challenging and exciting. I learned a lot. I learned how to pronounce some English sounds correctly. I would like to play such challenging games and activities again."

5. DISCUSSION

Research question 1 was intended to examine the effect of playing *Spaceteam ESL* on English pronunciation among a sample of Iranian EFL learners. The results provided supporting evidence for the effectiveness of game applications in improving EFL learners' pronunciation accuracy. Data analysis in the quantitative part showed that using the *Spaceteam ESL* application had an effect on the accurate pronunciation of English sounds, such as [w], [θ], [ð], [uː], [ʊ], [iː], and [ɪ]. One justification for pronunciation development is peer interaction and more oral output by the students in the experimental group. As the results in the qualitative part verified this issue, social feature of the game allowed the high school participants in the experimental group to speak with each other more and assisted them to produce more oral output, resulting in more pronunciation improvement. These high school EFL participants could practice their pronunciation in an interactive way and receive feedback from their classmates based on their success or failure in the gameplay. This justification can be in harmony with the prediction of Swain's (2000) hypothesis, claiming that learners learn through oral practice and output in their own L2. The students could correct their pronunciation during the gameplay. The game had pedagogical elements to entertain and train the students through different modes with regard to their pronunciation. The results about the effectiveness of the *Spaceteam ESL* game are partially in congruent with the findings of the prior research. Reinders and Wattana (2015) examined the effect of a digital game used in L2 education, namely, *MMORPG*, on interaction of 30 Thai learners of English. They concluded that compared with interaction in class, the online gameplay resulted in a larger increase in English interaction. Grimshaw and Cardoso (2018) also reported that playing the *Spaceteam ESL* game improved Canadian ESL learners' oral fluency and willingness to communicate.

Additionally, playing *Spaceteam ESL* game made the participants in the experimental group to repeat the English sounds and create a link between ear (listening) and tongue (pronunciation). Thus, the other justification for the effectiveness of the game in pronunciation development was a meaningful repetition feature demonstrated in the gameplay. As Turgut & Irgin, (2009) maintain, the natural repetition in learning games permits language learners to be persistently exposed to the target language, providing more opportunities for language acquisition to occur. Besides, attention should be paid to such aspects as native-like accent observed the *Spaceteam ESL* game. The participants pointed out features such as accent that stimulated them to repeat the target words and

facilitated the process of pronunciation acquisition/learning. The first phase of the game provided a great amount of aural input for the participants. The players could press one button and listen to the vocabulary and press another button and imitate the target vocabulary several times. As some learners contended, the input produced by the application was heavily accented native-like, which helped them improve their pronunciation.

As to research question 2, the qualitative data analysis suggested that the interviewees accentuated the fun and attractive atmosphere of the classroom. Reportedly, the attractive, fun, and pleasant atmosphere of the classroom encouraged them to pursue their activities and finish the class without feeling of discomfort, whereas such features were less observed in the traditional class. Gameplay used in the class was a novel experience for them and made them interested in pronunciation development. These aspects are also highlighted in Cho and Castaneda's (2019) study. They reported that the *Conjugation Nation* game was effective and considerably improved Spanish students' satisfaction after participating in game-like activities. They also pointed out the importance of attractiveness and fun nature of the class in which students played the *game*. Presumably, playing the game made the high school students in the experiment group of the study engaged, motivated and active in learning English pronunciation. This issue can correspond to the benefits of using technology in L2 learning. With the game-based instruction, their classroom became more student-centered, resulting in developing learner-centered learning.

The results revealed the role of games in reducing levels of anxiety and stress in foreign language pronunciation development. Teamwork, peer interaction, and fun atmosphere created as the result of using the game contributed to the low level of anxiety among the participants in the experimental group and, accordingly, facilitated pronunciation development. As Grimshaw and Cardoso (2018) maintain, the fun aspect of mobile gaming may aid learners to have a low level of anxiety, which is important for L2 pronunciation acquisition. The results also stressed elements such as challenges. That is to say, the game offered different difficulty levels and time limitation. The game-based instruction was a challenge for many of them. Conflict and challenge are primary elements of games that evoke players' arousal and make them feel creative, motivated, and exited (Prensky, 2001).

6. CONCLUSION AND IMPLICATIONS

The main aim of the study was to look into the possible effect of game-based instruction on the English pronunciation accuracy of high school students at low-intermediate level, and explore the motivational perception towards the game. To this end, this study used the *Spaceteam ESL* application. The results showed that the game-based instruction was effective and improved the EFL learners' pronunciation more than the traditional instructional method. By using the *Spaceteam ESL* game, the learners could create a link between ear (listening) and tongue (pronunciation). Using the *Spaceteam ESL* application helped the high school EFL students practiced the pronunciation of English sounds with a driving force. In addition, the results of the interviews revealed the applicability and attractiveness of this game application in increasing EFL learners' motivation. In the light of the emerging themes, it was shown that using such games as *Spaceteam ESL* could create a fun atmosphere in the classroom, lead to social interaction with peers, encourage EFL students to better their pronunciation, facilitate the process of English pronunciation acquisition, reduce anxiety level and stress, and stimulate them to accomplish the challenging game task.

This current study highlights the role of technology in L2 learning and contributes to broadening our understanding and consideration of motivational perceptions and serious engagement in MALL. The results can help L2 researchers extend gamification research to other aspects of L2 learning. The above findings can also have implications for L2 teachers. Based on the findings,

pedagogical games can offer advantages in L2 classes. L2 teachers can use games such as the one used in the present study as a supplement to other instructional materials in their L2 pronunciation programs. The results suggest that teachers can integrate the game with other audio and visual instructional materials to develop L2 pronunciation. In this way, they can support their students' active participation in the process of learning pronunciation in the target language in the classroom. Additionally, game-based instruction can stimulate L2 learners to enjoy the demanding pronunciation tasks in the process of learning L2 pronunciation. The findings imply that L2 teachers can employ gameplay to make their students experience a more pleasant and entertaining atmosphere in the L2 classroom. It is thus recommended that curriculum developers, materials writers, and course designers pay attention to the role of gameplay, game-like activities, and mobile-based activities in L2 teaching/learning and developing corresponding materials and curriculums.

Like many other studies, this study is not free from limitations. Sampling size is the first limitation. The number of participants was not large in order to make strong generalizations to the population of EFL learners. The second limitation is the level of the EFL learners which makes the generalizability to the other learners at a different proficiency level difficult. Furthermore, no randomization was possible in the study and the participants were from intact classes; this issue might also limit the generalizability of the data. Third, the study was implemented with a limited number of sessions. The length of the course may affect the findings. Thus, the results need to be interpreted cautiously. Finally, this study focused on the segmental aspect of pronunciation. Future researchers can validate the above findings by examining the role of game-based instruction with regard to different aspects of L2 pronunciation, utilizing different mobile applications over a long period of time and with learners at different L2 proficiency levels.

Declaration of Conflicting Interests
There is no conflict of interest to declare.

References
Alavi Moghaddam, S. B., Kheirabadi, R., Rahimi, M., & Davari, H. (2020). *English for schools (12th grade)- vision 3*. Tehran: Planning and Research Organization of the Ministry of Education.

Al-Zahrani, A. S. (2015). *Smartphones wandering at the MALL: A case study investigating the use of smartphones on English oral learning skills in a collaborative mobile-assisted language learning environment*. Unpublished doctoral dissertation, University of Northern Illinois.

Boo, Z., Dörnyei, Z., & Ryan, S. (2015). L2 motivation research: Understanding a publication surge and changing landscape. *System, 55*, 145-157. https://doi.org/10.1016/j.system.2015.09.001

Celce-Murcia, M., Goodwin, J., & Brinton, D. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.

Chang, M. (2005). Applying self-regulated learning strategies in a web-based instruction: An investigation of motivation perception. *Computer Assisted Language Learning, 18*, 217-230. https://doi.org/10.1080/09588220500123495

Chinnery, G. M. (2006). Going to the MALL: Mobile assisted language learning. *Language Learning & Technology, 10*(1), 9-16.

Cho, M. H., & Castaneda, D. A. (2019). Motivational and affective engagement in learning Spanish with a mobile application. *System, 81*, 90-99. https://doi.org/10.1016/j.system.2019.01.004

Coyne, R. (2003). Mindless repetition: Learning from computer games. *Design Studies, 24*, 199-212. https://doi.org/10.1016/S0142-694X(02)00067-0

Deng, J., Holtry, A., Howden-Weaver, L., Nessim, L., Nicholas, B., Nickle, K., Pannekoek, C., Stephan, M., & Sun, M. (2009). *English pronunciation research: The neglected orphan of second language acquisition studies?* Edmonton, AB: Prairie Metropolis Centre.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A search-based approach. *TESOL Quarterly, 39*, 379-397. https://doi.org/10.2307/3588533

Dörnyei, Z. (2001). *Teaching and researching motivation*. New York: Longman.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Harlow, England: Pearson Education.

Ebrahimzadeh, M., & Alavi, S. (2016). Motivating EFL students: E-learning enjoyment as a predictor of vocabulary learning through digital video games. *Cogent Education, 3*(1), 1-13. https://doi.org/10.1080/2331186X.2016.1255400.

Ena, T. O. (2003). Developing interactive computer-based language games using Microsoft PowerPoint. Malang: Graduate School of Malang University.

Fouz-González, J. (2018). Podcast-based pronunciation training: Enhancing FL learners' perceptions and production of fossilized segmental features. *ReCALL, 31*(2), 150-169. https://doi.org/10.1017/S0958344018000174

Fouz-González, J. (2020). Using apps for pronunciation training: An empirical evaluation of the English file pronunciation app. *Language Learning & Technology, 24*(1), 62-85.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Rowley, Massachusetts: Newbury House.

Ghapanchi, Z., Khajavy, G., & Asadpour, S. (2011). L2 motivation and personality as predictors of the second language proficiency: Role of the big five traits and L2 motivational self-system. *Canadian Social Science, 7*(6), 148-155.

Girard, C., Ecalle, J., & Magnant, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning, 29*, 207-219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Grimshaw, J., & Cardoso, W. (2018). Activate space rats! Fluency development in a mobile game-assisted environment. *Language Learning & Technology, 22*(3), 159-175.

Haghighi, M., & Rahimy, R. (2017). The effect of L2 minimal pairs practice on Iranian intermediate EFL learners' pronunciation accuracy. *International Journal of Research in English Education, 8*, 42-48.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford University Press.

Hismanoglu, M. (2006). Current perspectives on pronunciation learning and teaching. *Journal of Language and Linguistic Studies, 2*(1), 101-110.

Hsu, L. (2016). An empirical examination of EFL learners' perceptual learning styles and acceptance of ASR-based computer-assisted pronunciation training. *Computer Assisted Language Learning, 29*(5), 881-900. https://doi.org/10.1080/09588221.2015.1069747

Huizenga, J., Admiraal, W., Akkerman, S., & ten Dam, G. (2009). Mobile game-based learning in secondary education: Engaging, motivation, and mobile learning in a mobile city game. *Journal of Computer Assisted Learning, 25*(4), 332-344. https://doi.org/10.1111/j.1365-2729.2009.00316.x

Joseph, S., & Uther, M. (2006). Mobile language learning with multimedia and multi-modal interfaces. In *IEEE International Workshop on Wireless and Mobile Technologies in Education* (pp. 16-17). Greece, Athens.

Kétyi, A. (2013). Using smartphones in language learning – A pilot study to turn CALL into MALL. Retrieved from https://www.researchgate.net/publication/259009520_Using_Smart_Phones_in_Language_Learning_-_A_Pilot_Study_to_Turn_CALL_into_MALLA_Pilot_Study_to_Turn_CALL_into_MALL](https://www.researchgate.net/publication/259009520_Using_Smart_Phones_in_Language_Learning-_A_Pilot_Study_to_T

Lee, C. C., Hao, Y., Lee, K. S., Sim, S. C., & Huang, C. C. (2019). Investigation of the effects of an online instant response system on students in a middle school of a rural area. *Computers in Human Behavior, 95*, 217-223. https://doi.org/10.1016/j.chb.2018.11.033

Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education, 55*, 630-643. https://doi.org/10.1016/j.compedu.2010.02.008

Miangah, T. M., & Nezarat, A. (2012). Mobile-assisted language learning. *International Journal of Distributed and Parallel Systems, 3*(1), 309-319.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481-520. https://doi.org/10.2307/3586981.

Nurhayati, D. A. W. (2015). Improving students' English pronunciation ability through Go Fish game and Maze game. *Dinamika Ilmu, 15*(2), 215-233. https://doi.org/10.21093/di.v15i2.151

Pallier, C., Angels, C., & Sebastián-Gallés, N. (2019). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445-449. https://doi.org/10.1111/1467-928007381

Partovi, T., & Razavi, M. R. (2019). The effect of game-based learning on academic achievement motivation of elementary school students. *Learning and Motivation, 68*, 1-9. https://doi.org/10.1016/j.lmot.2019.101592

Pennington, M. C. (1996). *Phonology in English language teaching*. New York, NY: Addison Wesley Longman Limited.

Prensky, M. (2001). *Digital game-based learning*. New York, NY: McGraw-Hill.

Reinders, H., & Wattana, S. (2014). Can I say something? The effects of digital game play on willingness to communicate. *Language Learning & Technology, 18*, 101-123.

Reinders, H., & Wattana, S. (2015). The effects of digital game play on second language interaction. *International Journal of Computer-Assisted Language Learning and Teaching, 5*(1), 1-21. https://doi.org/10.4018/IJCALLT.2015010101

Salaberry, M. (2001). The use of technology for second language learning and teaching: A retrospective. *Modern Language Journal, 85*(1), 39-56. https://doi.org/10.1111/0026-7902.00096

Sandberg, J., Maris, M., & de Geus, K. (2011). Mobile English learning: An evidence-based study with fifth graders. *Computers & Education, 57*, 1334-1347. https://doi.org/10.1016/j.compedu.2011.01.015

Schwabe, G., & Göth, C. (2005). Mobile learning with a mobile game: Design and motivational effects. *Journal of Computer Assisted Learning, 21*, 204-216. https://doi.org/10.1111/j.1365-2729.2005.00128.x

Stockwell, G., & Hubbard, P. (2013). Some emerging principles for mobile-assisted language learning. Monterey, CA: The International Research Foundation for English Language Education.

Su, C. H., & Cheng, C. H. (2015). A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning, 31*(3), 268-286. https://doi.org/10.1111/jcal.12087

Sudarmilah, E., Irsyadi, F. Y. A., Purworini, D., Fatmawati, A., Haryanti, Y., Santoso, B., Bakhtiar, D. N., & Ustia, N. (2020). Improving knowledge about Indonesian culture with augmented reality gamification. Retrieved from https://iopscience.iop.org/article/10.1088/1757-899X/830/3/032024](https://iopscience.iop.org/article/10.1088/1757-899X/830/3

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford University Press.

Turgut, Y., & Irgin, P. (2009). Students' experiences of computer games and ELL: From a phenomenological perspective. *Proceeding of the 10th METU ELT Convention, 22*, 1-7.

Ushioda, E. (2013). Motivation matters in mobile language: A brief commentary. *Language Learning and Technology, 17*(3), 1-5.

Viberg, O., & Grönlund, A. (2012). Mobile assisted language learning: A literature review. Retrieved from https://www.researchgate.net/publication/277832223_Mobile_Assisted_Language_Learning_A_Literature_Review](https://www.researchgate.net/publication/277832223_Mobile_Assisted_Language

Wastiau, P., Kearney, C., & Van den W. (2009). *How are digital games used in school* Belgium: European Schoolnet.

Zheng, D., Young, M. F., Brewer, R. A., & Wagner, M. M. (2009). Attitude and self-efficacy change: English language learning in virtual worlds. *CALICO Journal, 27*, 205-231.

Language Proficiency and EAP Students' Writing Quality: Contributions of Explicit and Implicit Genre-Based Instruction

Sedigheh Karimpour[1,*] iD

[1] Mazandaran University of Medical Sciences, Sari, Iran

Abstract
Genre-based pedagogy provides an effective resource for writing instructors and practitioners in order to assist their students to produce particular texts (Bhatia, 2004). Regarding the benefits and the uses of genre-based pedagogy in second and foreign language acquisition, ongoing debates among the researchers of the field have been made in response to pedagogical and implicational practices associated with genre in various educational settings. Due to the importance of providing writing instructions for the students of medicine and the lack of literature in analyzing the impact of different treatments of genre-based instruction especially among students with different language proficiency levels, the present study tried to address this gap by taking a genre-based approach and using writing practice, to examine the impacts of explicit and implicit genre-based instructional techniques on medical students' writing quality with different language proficiency levels. The initial findings of the study suggested that providing genre-based instruction both explicitly and implicitly can intuitively and effectively assist learners in providing more qualified writings. Moreover, the in-depth findings of this study supported the proposal that higher proficient second language writers get more advantage of both treatments of genre-based instruction than lower proficient second language writers especially when explicit genre-based instruction is provided.

1. INTRODUCTION

Over the past few decades, English for specific purposes (ESP) researchers revolutionized the concept of genre, challenged the proposal that different genres are different categorizations of text types, and suggested that genre connects various texts to various social actions (Hyland, 2018). The conception of genres, as various rhetorical ways of interaction within recurring settings, has had a paramount impact on the study and teaching of writing. In this regard, the analytical and pedagogical implications of genre, which considered it as a significant variable in writing acquisition, have been explored by researchers (Bawarshi & Reiff, 2010).

[*] **Corresponding Author**: Sedigheh Karimpour, Department of Foreign Languages Mazandaran University of Medical Sciences Sari, Iran. **Email**: sedighehkarimpour1367@yahoo.com

Genre has been defined and used differently in different academic contexts. In various disciplines, researchers defined genre as a way of organizing various types of texts and other cultural events (Hyland, 2009, 2018). According to Hyland (2009), a genre-based approach to writing delves into the structural organization of texts in order to understand how writing works as communication. According to genre-based approaches, texts are written as a response to a particular communicative context. Therefore, the role of writing is seen as a mediator between institutions and cultures in which it occurs. In other words, texts convey the writers' purposes and expectations about how information should be structured (Hyland, 2009).

Regarding the impacts and benefits of genre-based writing instruction in second and foreign language acquisition, there have been ongoing debates among researchers of the field in pedagogical practices associated with genre-based instruction in various educational settings (Khodabandeh et al., 2013). While some researchers (e.g., Hammond & Derewianka, 2001) doubted the effectiveness of explicit genre-based instruction, others (e.g., Freedman, 1993; Bhatia, 2001) emphasized the effectiveness of explicit instruction of genres (e.g., Skillen & Trivett, 2001). Despite the benefits of genre-based instruction in second language writing (Bawarshi & Reiff, 2010; Hyland, 2008), there have been few experimental studies on the impacts of explicit and implicit genre-based instruction on writing quality of the students with different language proficiency levels.

In light of this understanding, the current study made an attempt to analyze and examine the writing quality of high and low proficient students of medicine who were exposed to explicit and implicit genre-based instruction. Therefore, one recurrent type of medical genre (medical case-report) was used. Medical case-reports serve as professional and educational narratives that provide feedback on clinical practice and offer a framework for sharing medical, scientific, or educational purposes (Florek & Dellavalle, 2016).

Due to the importance of providing writing instructions for academic students (Teeter, 2016; Khakpour & Shahsavar, 2019) and the lack of research on analyzing the impacts of explicit and implicit genre-based instruction on high and low proficient students of medicine, the present study attempted to explore the effects of such treatments on writing quality of students with different language proficiency levels. The study, thus, makes a contribution to the body of knowledge on genre-based instruction and helps ESP teachers and students gain awareness of the role of genre in their disciplines, particularly here medical science students. Therefore, the present study attempted to respond two research questions:

1) Which method of genre-based instruction (explicit or implicit) is more effective in developing EAP students' writing quality? 2) Is there any statistically significant difference in high and low proficiency EAP students' writing quality in terms of genre?

2. LITERATURE REVIEW

Genre-based Approaches to Writing Instruction

During the last decade, researchers in second and foreign language acquisition have given increasing attention to genre and the application of genre-based instruction in language teaching (Shaw, 2016). Genre-based instruction (GBI) offered opportunities for second language teachers to provide students with relevant and especially supportive instruction in enabling the students to follow standardized approaches while composing texts (Hyland, 2018). GBI enables L2 teachers to manage their courses in a way that students will need to write in future occupational, academic or social contexts that they may encounter (Hyland, 2018).

Genre-based instruction assumes that through acquiring genre knowledge in writing, students will learn how to analyze a genre in detail, which is suitable to various requirements of the

discourse community. It is believed that the lexical–grammatical patterns of the texts should be known and mastered by the students, and the teacher's task is to enhance the students' awareness of the target genres (Worden, 2019).

Explicit and Implicit Genre-Based Instruction

Despite the benefits of genre-based instruction in second language writing (Hyland, 2008; Bawarshi & Reiff, 2010), the number of studies with the focus on genre-based instructional treatments is low. Hyon (2001) investigated the impact of explicit teaching of genre on students' mastery of rhetorical features. His findings revealed that explicit genre-based instruction facilitates different aspects of learners' second language writing (Hyon, 2001). The findings of the study by Pang (2002) indicated that contextual awareness and textual analysis as two testaments of genre-based instruction enhanced students' employment of move structure in genre.

Abbuhl1 (2011) analyzed the impacts of explicit and implicit genre-based instruction on university students' writing development and revealed that the participants improved their writings after receiving explicit genre-based instruction. Khodabandeh et al. (2013) investigated the impacts of three treatments of genre-based pedagogy on students' quality of writing argumentative essays. Their findings revealed that explicit teaching was more effective in assisting participants to provide accurate text essays.

Despite the effectiveness of genre-based instruction in different writing classes, few studies have been done so far with the focus on the impacts of explicit or the implicit genre-based instruction on students' writings in different language proficiency levels. Due to the lack of research on the efficacy of genre-based pedagogy (Tardy, 2006), the present study attempted to address this gap in the realm of medical sciences and in relation to proficiency level as a mediating factor.

Language Proficiency and Second Language Writing Ability

Researchers have so far investigated factors affecting second language learners' texts. Factors such as learners' second language proficiency, second language writing strategies and first language writing ability are contributive to the way students structure their texts (Hyland, 2018). Investigation of the impact of L2 proficiency on writing skill has yielded mixed results (Sasaki & Hirose, 1996). Previous research has highlighted that learners' linguistic proficiency does not influence their writing ability and writing development (e.g., Raimes, 1985; Zamel, 1982). These studies maintained that learners' linguistic competence is not the determining and important factor influencing students' writing quality. On the other hand, some studies have suggested learners' second language proficiency as an important factor in the quality of writing. Cumming (1989), for example, revealed that students' L2 proficiency was a distinct factor that affected their writing quality. Pennington and So (1993) found that among several factors investigated, students' L2 proficiency was a determining factor to identify good writers.

Raimes (1987) examined ESL student writers at different levels of instruction to describe their writing strategies. The findings of his study revealed that language proficiency had almost no observable effect on composing strategies. Kobayashi and Rinnert's (1990) study argued that high proficient students tended to make more errors in their writing tasks than low proficient students. Moreover, Stevens et al. (2000) analyzed the performance of students on two different tests, namely a standardized achievement test and a language proficiency test. Their findings showed that there was a limited connection between the languages of the two tests. On the other hand, Wellington (2000) found a strong positive relationship between students' performance on an achievement test and their language proficiency level. Vazquez, Vazquez, and Lopez (2014) examined the correspondence between learners' language proficiency and their academic writing

attainment. The findings of the study revealed that the connection between English proficiency and standardized achievement scores was significant.

Although most of previous studies provided insight into the relationship between language proficiency as a factor contributing to the writing skill, few studies have so far investigated the impact of genre-based writing instructional treatments on EAP students with different language proficiency levels. With these limitations in mind, the present study examined Iranian Medical university students with appropriate controls for their educational and cultural backgrounds.

English for Medical Purposes

Researchers believe that English for medical purposes as a subsection of ESP aimed to provide the professional and academic needs of medical students, particularly in relation to writing (Paltridge, 2004). English for medical purposes provides special kinds of lexical repertoire and language skills that medical students and medical practitioners are likely to require in the workplace (Antic, 2007). English for medical purposes assists students by providing exposure to authentic texts and materials, and affordsthem the specific rhetorical and discourse features in order to facilitate comprehension of the specialized texts which they may encounter in their future workplace (Porcaro, 2013).

From the field of professional medical writing, one type of text, which is used as part of the everyday routine in hospitals and other health-care facilities and one of the genre types, which has recently received a close attention, is the medical case-report (Helan, 2012). Case-reports provide opportunities for medical practitioners and students through which they can communicate the evaluation and diagnosis of different clinical conditions and diseases (Felorek & Dellavalle, 2016). Case-report has always been an important tool for medical education and it presents challenging medical cases to medical students and medical practitioners (Felorek & Dellavalle, 2016). Concerning the importance of medical case-report as a frequent medical written text type for the students of medicine (Nissen & Wynn, 2014; Yasuda, 2011), the present study used this genre in order to provide genre-based written instruction. Thus, the study aims to examine differences in high and low proficiency medical students and variations in the genre-specific texts they have composed.

3. METHOD

Design

This study used a quasi-experimental, counter-balanced design. The method used in the present study has been previously used by (Alavi et al., 2021). The scores from the pre and post essay's scores were used and analyzed in order to understand whether the students improved their writing after receiving interventions. The students' writings were analyzed based on the rhetorical structure model of case report genre adapted by Hung et.al. (2012). The present study adopted Hyland's (2008) teaching–learning cycle model for teaching the selected academic genre. This model consists of four major stages of modeling, joint construction, independent construction, and comparing for this model.

Table 1: Counter-Balanced Design of Study

Pretest	Intervention	Posttest
Pretest	Explicit First	Posttest
Pretest	Implicit First	Posttest
Pretest	Implicit Second	Posttest
Pretest	Explicit Second	Posttest
Pretest	Control	Posttest

As it is shown in Table 1, in the first step, the data was collected from the participants using the pre-test. The first instrument used by the researchers was a general language proficiency test and the second instrument was an academic writing test. A pre-test was administered to measure the writing quality of the students before the intervention and then after the intervention, post-tests were administered to measure the impact of each of the interventions on students' writing quality.

Participants

A group of 150 medical students were selected conveniently from Mazandaran University of Medical Sciences. The students of the present study included both male and female students. The average age of the student participants of the study was 20 years old. The participants of the study did not receive any genre-based instruction before the study after consulting them about this issue.

Instrumentation

The first instrument was a general language proficiency test. In this study, to determine the students' language proficiency level, the instructor administered a modified version of IELTS Proficiency Test to all the participants. The IELTS exam test, as a reliable mean of grading students at all levels from elementary upwards, provides a consistent record of predictive validity in respect of examination entry (Freimuth, 2009). Due to the effectiveness of IELTS in assessing language learners' proficiency levels, at the beginning of the semester, the test was administered to the participants of the study in order to determine their level of language proficiency. The second instrument was an academic writing test. The three academic writing tests of the present study were as follows: a pre-test and two post-tests at the end of each intervention. In order to find out whether the students improved their academic writing quality, the three academic writing tests were administered.

Procedure

The teaching materials and activities which were used in the classes were prepared in line with the objectives of each group. Each group of the study had its own outline and they all shared the same case report examples. The order of the sessions was according to the four stages of the teaching-learning cycle model (Hyland, 2008). In order to apply explicit genre-based instruction, the teacher provided the models of medical case report genre and they practiced them in their class. Various structural and rhetorical features of case report genre were marked and underlined in order to highlight the rhetorical moves of this specific genre. In order to apply implicit genre-based instruction, the participants were exposed to the same model genres, yet the major moves of the genre were not highlighted for the students and they were not named by the instructor. In the instructional sessions, sample case reports were given to the students. Similar to the genre-based writing cycle (Hyland, 2008) which were practiced with the first and second experimental groups, the subjects of control group received the same treatment except the modeling stage. This group did not receive the researcher's treatments and the students were taught based on the outline designed by the department. The data from the students' writing were analyzed via one-way analysis of variance (ANOVA).

4. RESULT AND DISCUSSION

Results

One-way analysis of variance was run to compare the three groups' means on pretest in order to prove that they were homogenous with regard to their writing quality prior to the administration of the treatment. The results of the Levene's test of homogeneity of variances ($F(2, 27) = 1.08$, $p = .353$) indicated that the three groups enjoyed homogenous variances on pretest. Based on the results of one-way ANOVA, (F (2, 27) = .596, p = 560, Partial eta squared = .042 representing a weak effect size), it can be inferred that differences between the three groups' means on pretest were not significant. Thus, it suggests that they were homogenous with respect to their writing quality prior to the main study. The participants took two posttests at two-time intervals during which the treatments of the two experimental groups were shifted, as displayed below.

A one-way ANOVA was run to compare the three groups' means on posttest of writing quality. Since the absolute values of the ratios were lower than 1.96, it can be claimed that the assumption of normality was retained. The findings showed that the first experimental group who received explicit genre-based instruction first and then received implicit genre-based instruction (M = 41.80) outperformed the second experimental group in improving their writing quality (M= 39.90).

The second research question explored whether the difference between high and low proficiency groups' means on posttest of writing quality was significant. It should be noted that the present sample was divided into two proficiency levels of high and low based on the median score of 5. As illustrated in Table 3, the participants' scores on the IELTS test had mean of 5.27, a median and mode of 5.

Table 2: Descriptive Statistics; Posttest by Groups

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
Control	10	12.90	3.479	1.100	10.41	15.39	8	19
1stExplicit	10	35.70	4.001	1.265	32.84	38.56	29	41
1stImplicit	10	26.10	7.156	2.263	20.98	31.22	16	37
2ndImplicit	10	40.80	3.706	1.172	39.15	44.45	36	45
2ndExplicit	10	38.90	4.483	1.418	35.69	42.11	31	45
Total	50	31.08	11.558	1.635	27.80	34.36	8	45

Table 3: Descriptive Statistics of IELTS

Valid	30
Missing	0
Mean	5.27
Median	5.00
Mode	5
Std. Deviation	.640

Table 4: Descriptive Statistics; Testing Normality of Posttest by Proficiency Levels

Proficiency	N	Skewness			Kurtosis		
	Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
Low	30	-.546	.427	-1.27	-1.317	.833	-1.58
High	20	-.875	.512	-1.72	-.720	.992	-.725

Table 5: Descriptive Statistics; Posttest of Writing by Proficiency Levels

	Group	N	Mean	Std. Deviation	Std. Error Mean
Posttest	High	20	35.75	10.213	2.284
	Low	30	27.97	11.503	2.100

Table 6: Independent Samples t-test; Posttest of Writing by Proficiency Levels

	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	1.207	.277	2.449	48	.018	7.783	3.178	1.393	14.174
Equal variances not assumed			2.509	44.073	.016	7.783	3.102	1.531	14.036

An independent-samples t-test was run to compare the high and low proficiency groups' means on posttest of writing quality because the assumption of normality, as displayed in Table 4, was retained. Since the absolute values of the ratios were lower than 1.96, it can be claimed that the assumption of normality was retained.

Based on the results portrayed in Table 5, it can be claimed that the high proficiency group (M = 35.75, SD = 10.21) had a higher mean than the low proficiency group (M = 27.97, SD = 11.50) on the posttest of writing.

The results of the independent t-test (t (48) = 2.44, p = .018, Cohen's d = .707 depicting a moderate to large effect size) (Table 6) indicated the high proficiency group significantly outperformed the control group on the posttest of writing. Thus it can be claimed that there is a statistically significant difference in medical students' writing quality in terms of genre across participants with different levels of language proficiencies.

It is worth noting that the assumption of homogeneity of variances was met (Levene's F = 1.20, p = .277). In this respect, the first row of Table 6, i.e. "Equal variances assumed" was reported.

5. DISCUSSION

The present study explored impacts of explicit and implicit genre-based writing instruction on medical student' writing quality with different levels of language proficiency. In order to decrease the risk of over-emphasizing product view toward writing, this study utilized teaching and learning cycle model (Hyland, 2008). By doing so, the impact of explicit and implicit genre-based instructional techniques was investigated among the students with different levels of language proficiency.

The first research question explored whether explicit or implicit genre-based instruction were effective in developing the writing quality of the participants. The analyses of the students' data yielded the existence of a significant difference in the impact of both of the treatments of genre-based instruction on the writing quality of the participants. Moreover, there were significant statistical differences among the two experimental groups and the control group on the post-tests of writing quality. Also, findings of the study confirmed that differences between the two experimental groups was not significant. The results of the present study revealed the potential of explicit genre-based writing which is in line with Swami (2008), Abbuhl (2011) and Khodabandeh et.al (2013) who reported the efficacy of explicit teaching methods. However, unlike the findings by Henry and Roseberry (1998) which indicated that it is not possible to gain genre awareness in the absence of explicit instruction (Khodabandeh et.al, 2013), the results of this study confirmed that the students benefited from implicit instruction and in the absence of explicit instruction they improved their writing quality. This finding shows that students can benefit from both types of GBI, yet explicit focus on genre makes the type of text more tangible for the students.

The second research question aimed to scrutinize if the differences between high and low proficiency students in case of writing quality was statistically significant. The analyses of the students' data showed that the high proficiency group had a higher mean than the low proficiency group on the posttest of writing. The findings of the study implied that the performance of high proficiency group significantly exceeded than the performance of the control group on the posttests.

Unlike the study by Kobayashi and Rinnert's (1990) in which higher-proficient students made more structural and rhetorical errors than lower-proficient students, the findings of the present study revealed that higher-proficient students tended to make less errors than lower-proficient students. Moreover, the results of this study is in line with those by Newman, Tremblay, Nichols, Neville, and Ullman (2012) and Vazquez, Vazquez and Lopez (2014) which revealed significant correlations between writing quality and language proficiency. These findings mean that as proficiency level increases, students not only pay focal attention to structural issues in composing effective essays, but they also engage in honing the wide range of skills needed for writing error-free texts.

6. CONCLUSIONS AND IMPLICATION

The findings of the study suggested that providing genre-based instruction both explicitly and also implicitly can intuitively and effectively help learners to learn main generic moves of a genre. Also applying either of these treatments may lead to a better use of the moves and hence more qualified writing. The findings of the study confirmed the proposal that providing either explicit or implicit genre-based instruction for higher proficient students leads to developments in their writing quality. In other words, combining explicit genre-based instruction with the implicit genre-based instruction was likely to result in better outcomes for higher proficient students rather than teaching through using one of these treatments on its own. In sum, it can be concluded that explication of genre-based instruction certainly had vast potential benefits to medical students in using genre moves in their case report essays. The findings of this study also supported the proposal that higher proficient L2 writers get more advantage of genre-based instruction than lower proficient L2 writers especially when genres are taught explicitly.

The present study has some implications that are worth noting. This research study raised a number of issues related to the teaching of writing for EAP instructors that they should take into consideration in their teaching of academic writing. It is suggested that the instructors analyze the teaching of writing from an ESP perspective with the attention on meaning construction at a discursive level. Learning about move structures of the genre may help learners to organize their

texts in terms of knowing what to write. Moreover, ESP instructors can make use of both the explicit and implicit treatments of genre-based instruction in order to help their students to improve their writing qualities and writing developments. Furthermore, this study drew on two proficiency levels as the basis for examining genre-based differences. Further research with more groups would provide a better picture of how proficiency level mediates genre-based texts students compose.

References

Abbuhl, R. (2011). Using models in writing instruction: A comparison with native and nonnative speakers of English. *SAGE Journals, 1*(3), 1-12. https://doi.org/10.1177/0033688211422901

Allwright, J., & Allwright, R. (1977). An approach to the teaching of medical English. In S. Holden (Ed.), *English for Specific Purposes* (pp. 58-62). London: Modern English Publications.

Anderson, M., & Anderson, K. (2003). *Text Type in English*. South Yarra: Macmillan Education Australia Pty Ltd.

Badger, R., & White, G. (2000). A process genre approach to teaching writing.

Bakhtin, M. (1981). *The Dialogic Imagination: Four Essays by M. M. Bakhtin*. Austin: University of Texas Press.

Bawarshi, A., & Reiff, M. (2010). *Genre, an Introduction to History and Theory, Research and Pedagogy*. Parlor Press.

Bhatia, V. (2001). The power and politics of genre. In A. Burns & C. Coffin (Eds.), *Analyzing English in a Global Context: A Reader* (pp. 65-77). London: Routledge.

Bhatia, V. (2004). *Worlds of Written Discourse: A Genre-Based View*. London: Continuum.

Brown, J. D., & Bailey, K. M. (1984). A categorical instrument for scoring second language writing skills. *Language Learning, 34*(4), 21-42. https://doi.org/10.1111/j.1467-1770.1984.tb00353.x

Bruce, I. (2010). Textual and discoursal resources used in the essay genre in sociology and English. *Journal of English for Academic Purposes, 9*, 153-166. https://doi.org/10.1016/j.jeap.2010.02.004.

Bruce, I. (2008). *Academic Writing and Genre: A Systematic Analysis*. London: Continuum.

Burke, K. (1974). *The Philosophy of Literary Form: Studies in Symbolic Action*. Berkeley: University of California Press.

Campbell, K. K., & Hall Jamieson, K. (1990). *Deeds Done in Words: Presidential Rhetoric and the Genres of Governance*. Chicago: University of Chicago Press.

Canagarajah, S. (1999). *Resisting Linguistic Imperialism in English Teaching*. Oxford: Oxford University Press.

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

Chapman, M. (1999). Situated, social, active: Rewriting genre in the elementary classroom. *Written Communication*, 469-490. https://doi.org/10.1177/0741088399016003001

Cheng, A. (2008). Analyzing genre exemplars in preparation for writing. *Applied Linguistics, 29*(1), 50-71. https://doi.org/10.1093/applin/amm021

Coe, R. (2002). The new rhetoric of genre: Writing political briefs. In A. M. Johns (Ed.), *Genre in the Classroom: Multiple Perspectives* (pp. 121-141). Mahwah, NJ: Lawrence Erlbaum Associates.

DeKeyser, R. (1995). Learning second language grammar rules. *Studies in Second Language Acquisition, 17*, 379-410. https://doi.org/10.1017/S027226310001422X

Devitt, A. J. (2006). First-year composition and antecedent genres. Presented at the Conference on College Composition and Communication, Chicago.

Dudley-Evans, T. (1998). An overview of ESP in the 1990s. The Japan Conference on English for Specific Purposes Proceedings. Aizuwakamatsu: Aizu University.

Ellis, R. (1993). The structure syllabus and second language acquisition. *TESOL Quarterly, 27*(1), 91-113. https://doi.org/10.2307/3586953

Florek, A., & Dellavalle, R. (2016). Case reports in medical education. *Journal of Medical Case Reports*. https://doi.org/10.1186/s13256-016-0744-4

Freedman, A. (1987). Development in story writing. *Applied Psycholinguistics, 8*, 153-165. https://doi.org/10.1017/S0142716400000143

Freedman, A. (1993). Show and tell? The role of explicit teaching in the learning of new genres. *Research in the Teaching of English, 27*, 222-251.

Gao, J. (2007). Teaching writing in Chinese universities: Finding an eclectic approach. *Asian EFL Journal, 20*(2).

Halliday, M. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.

Halliday, M., & Martin, J. (1993). *Writing Science: Literacy and Discursive Power*. Bristol, PA: Falmer Press.

Hammond, J., & Derewianka, B. (2001). Genre. In R. Carter & D. Nunan (Eds.), *The Cambridge Guide to Teaching of English to Speakers of Other Languages* (pp. 186-193). New York: Cambridge University Press.

Helan, R. (2012). Analysis of published medical case reports: Genre-based study. Unpublished dissertation, Masaryk University, Brno.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*(3), 148-164. https://doi.org/10.1016/j.jslw.2007.07.005

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching, 41*(4), 543-562. https://doi.org/10.1017/S0261444808005231.

Hyon, S. (2002). Genre and ESL reading: A classroom study. In A. M. Johns (Ed.), *Genre in the Classroom: Multiple Perspectives* (pp. 121-141). Mahwah, NJ: Lawrence Erlbaum.

Khakpour, N., & Shahsavar, Z. (2019). Comparing the use of different prewriting strategies on medical students' L2 writing. *Asian ESP Journal*, 228-255. https://doi.org/10.1007/s12144-019-00484-4

Macken-Horarik, M. (2002). Something to shoot for. In A. M. Johns (Ed.), *Genres in the Classroom: Applying Theory and Research into Practice* (pp. 17-42).

Mackey, A., Gass, S., & Ross, L. (2005). Task-based interactions in classroom and laboratory settings. https://doi.org/10.1111/j.1467-1770.2005.00318.x

Mazloomi, S. (2011). The impact of genre-awareness raising in EFL essay writing classes. *IPEDR, 26*, 106-111.

Nissen, T., & Wynn, R. (2014). The clinical case report. *BioMed Central*. https://doi.org/10.1186/1752-1947-8-390.

Paltridge, B. (2001). *Genre and the Language Learning Classroom*. Ann Arbor: University of Michigan Press.

Paltridge, B. (2004). Approaches to teaching second language writing. 17th Educational Conference Adelaide.

Paltridge, B., & Starfield, S. (2013). *The Handbook of English for Specific Purposes*. Boston: Wiley-Blackwell. https://doi.org/10.4000/asp.3806.

Skillen, J., & Trivett, N. (2001). Explicit teaching and assessment of genre conventions in university education. *Academic Exchange Quarterly, 5*, 113-118.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Swales, J. M., & Lindermann, S. (2002). Teaching the literature review to international graduate students. In A. M. Johns (Ed.), *Genre in the Classroom: Multiple Perspectives* (pp. 105-119). Mahwah, NJ: Lawrence Erlbaum.

Yasuda, S. (2011). Genre-based tasks in foreign language writing. *Journal of Second Language Writing, 20*, 111-133. https://doi.org/10.1016/j.jslw.2010.12.001

Yates, J. (1989). *Control Through Communication: The Rise of System in American Management*. Baltimore: Johns Hopkins University Press.

Yates, J., & Orlikowski, W. (1992). Genres of organizational communication. *Academy of Management Review, 17*, 299-326. https://doi.org/10.5465/amr.1992.4281175

Yates, J., & Orlikowski, W. (2002). Genre systems: Structuring interaction. *Journal of Business Communication, 39*, 13-35. https://doi.org/10.1177/002194360203900102

Yeh, S. (1998). Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English, 33*, 49-83.