International Journal of Foreign Language Teaching and Research
ISSN: 2322-3898-http://jfl.iaun.ac.ir/journal/about

Please cite this paper as follows:
Khaki, M., Heidari Tabrizi, H., & Shafiee, S. (2023). Effects of Immersive Learning and Communicative Language Teaching on the Oral Complexity, Accuracy, and Fluency of Iranian EFL Learners with Expressive Language Disorder. *International Journal of Foreign Language Teaching and Research, 11* (46), 63-81. http://doi.org/10.30495/JFL.2023.703378

Research Paper

Effects of Immersive Learning and Communicative Language Teaching on the Oral Complexity, Accuracy, and Fluency of Iranian EFL Learners with Expressive Language Disorder

Mojgan Khaki[1], Hossein Heidari Tabrizi[2] *, Sajad Shafiee[3]
[1]Ph.D. Candidate in TEFL, Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
Mojgan.khaki96@gmail.com
[2]Associate Professor, Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
Heidaritabrizi@gmail.com
[3]Assistant Professor, English Department, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran

Received: July 03, 2023 **Accepted: August 18, 2023**

Abstract
The CTML served as the primary source of motivation for this investigation, which explored the effects of Immersive Learning (IL) and Communicative Language Teaching on the Oral Complexity, Accuracy, and Fluency (CAF) of Iranian EFL learners with Expressive Language Disorder (ELD). Using a quasi-Solomon four-group design, the participants, who were all Iranian pre-intermediate EFL learners with mild to moderate levels of ELD, were separated into experimental and control groups. There were a total of 36 participants in this study. The intervention consisted of virtual reality (VR) authoring eighty speech exercises in the form of learning chunks for twenty sessions lasting seventy minutes each. In the groups that served as controls, the content was presented using a method known as Communicative Language Teaching (CLT). Both the pre-test and the post-test were conducted orally using the CAF test. In terms of oral complexity (syntactic complexity and diversity), oral accuracy (error-free sentences and accurate verb forms), and oral fluency (number of vocally generated syllables), the findings revealed that individuals who received IL fared better than their counterparts in the CLT group. The findings may have important implications for teaching English as a foreign language and for rehabilitative education.

Keywords: *EFL learners; Expressive Language Disorder; Oral accuracy; Oral complexity; Oral fluency*

تأثیر یادگیری فراگیر و آموزش ارتباطی بر پیچیدگی، دقت و روانی زبان آموزان ایرانی مبتلا به اختلال بیانی زبان
CTMLبه عنوان منبع اصلی انگیزه برای این تحقیق عمل کرد که اثرات یادگیری فراگیر (IL) و آموزش زبان ارتباطی بر پیچیدگی، دقت و روانی زبان (CAF) زبان آموزان ایرانی زبان انگلیسی مبتلا به اختلال زبان بیانی (ELD) را بررسی کرد. با استفاده از طرح چهار گروهی شبه سلیمان، شرکت کنندگان که همگی زبان آموزان ایرانی پیش از متوسط زبان انگلیسی با سطوح خفیف تا متوسط ELDبودند، به دو گروه آزمایش و کنترل تقسیم شدند. در مجموع ۳۶ شرکت کننده در این مطالعه حضور داشتند. مداخله شامل واقعیت مجازی (VR) بود که هشتاد تمرین گفتاری را در قالب تکه های یادگیری به مدت بیست جلسه به مدت هفتاد دقیقه نوشت. در گروههایی که به عنوان کنترل عمل می‌کردند، محتوا با استفاده از روشی به نام آموزش زبان ارتباطی (CLT) ارائه شد. هر دو پیش آزمون و پس آزمون به صورت شفاهی با استفاده از آزمون CAF انجام شد. از نظر پیچیدگی شفاهی (پیچیدگی و تنوع نحوی)، دقت شفاهی (جملات بدون خطا و اشکال افعال دقیق)، و روانی شفاهی (تعداد هجاهای صوتی تولید شده)، یافته‌ها نشان داد که افرادی که IL دریافت کردند بهتر از همتایان خود عمل کردند. گروه CLT این یافته ها ممکن است پیامدهای مهمی برای آموزش زبان انگلیسی به عنوان یک زبان خارجی و آموزش توانبخشی داشته باشد.
کلمات کلیدی: زبان آموزان زبان انگلیسی. اختلال زبان بیانی؛ دقت شفاهی؛ پیچیدگی دهان؛ روانی شفاهی

Introduction

A wide variety of instructional strategies, educational tools, interactive platforms, and telerehabilitation services have been offered with the purpose of being of assistance to those who suffer from learning disorders (LDs). At the same time, the vast majority of them took use of a sizable number of benefits, including advice from peers to peers, the production of content that was specifically customized to their requirements, user-friendly environments, and remote access. On the other hand, the item that is missing from this list is one that is compelling enough to arouse the interest of people with a variety of LDs. This problem receives a lot of attention when it comes to students or learners who are members of Generation Y (millennials) or Generation Z and have certain expectations about the solutions they employ in their academic careers. These expectations are noticeably different from those of their contemporaries who are members of Generation X (Cockrell, 2021). According to the Human, Activity, and Assistive Technology (HAAT) model in Assistive Technology (AT), the factors belonging to the three mentioned domains that are most important are those associated with the human aspect, including the motivation of this overlooked community of learners (Cook & Polgar, 2014). Triggering individuals with LDs to use AT depends on the factors belonging to the three mentioned domains, the most important ones being those associated with the human aspect.

Immersive Learning (IL) is the experience of employing Virtual Reality (VR) technology in the form of various solutions, such as non-immersive, semi-immersive, and fully-immersive platforms, with the goal of offering a non-biased, genuine, and engaging learning experience (Frehlich, 2020). Immersive Learning may take the shape of many solutions, including non-immersive, semi-immersive, and fully-immersive environments. The experience of using 2D images and videos without having any direct interaction with the environment, such as in video games, is referred to as non-immersive virtual reality (VR). On the other hand, fully-immersive virtual reality (VR) is complete engagement in a 3D environment similar to the real world in which users are able to move around and directly interact with the elements of the environment by using Virtual Reality Head-Mounted-Displays (VR-HMDs) and joysticks. The term "semi-immersive virtual reality" (VR) refers to the experience of utilizing restricted 3D pictures or videos in which the users are not able to interact with the environment directly but rather do so via indirect connections. One example of this kind of interaction is the experience of using simulators. According to Greengard (2019), the most realistic virtual reality platform is one that is completely immersive and offers a variety of sensory inputs delivered through a variety of sensory channels.

According to the Cognitive Theory of Multimedia Learning (CTML), the effectiveness of the learning process will increase in proportion to the number of sensory channels that are used throughout it (Mayer, 2009). This is because the increased cognitive load will be dispersed over a greater number of channels. In the LDs domains, the great efficacy potential of IL as well as its dynamic and engaging environment, which is considerably motivating for users of Generations Y and Z, becomes of crucial importance (Sabry, 2022). Speech or language problems, such as apraxia, dysarthria, receptive language disorder (RLD), and expressive language disorder (ELD), are examples of one of the most prevalent categories of learning disabilities (LDs). ELD is one of the most common speech disorders. It is not caused by articulation deficiencies and primarily affects verbal communication skills to a moderate degree. This is in contrast to other difficulties, such as articulation defects or an intelligence quotient (IQ) deficit, which not only affect verbal communication skills but also understanding a language and the cognitive aspects of speaking (Westwood, 2017).

When it comes to verbal communication abilities, oral Complexity, Accuracy, and Fluency (CAF) plays an important part in the process of meaning transmission. The concept of oral CAF

(Housen et al., 2012) is a legitimate and reliable technique to assess the oral or speaking abilities among English-speaking persons, whether in their mother tongue or in a foreign language. This applies whether the individual is speaking their mother tongue or a foreign language. Oral complexity in this construct refers to syntactic complexity as well as syntactic variation. Oral accuracy refers to error-free sentences as well as proper verb forms. Oral fluency refers to the amount of syllables that are generated verbally. Although a number of studies have investigated the affordance issues in relation to improving oral communication in EFL classes (Yudintseva, 2023), the modality and task complexity on oral performance in VR environment (York, 2019), and the pattern in verbal communication in high immersion VR (Gruber & Kaplan-Rakowski, 2022), there is still a large gap in considering the integration of therapeutic speech programs objectives which were based on telerehabilitation.

Although the majority of studies came to the subjective conclusion that the use of virtual reality technology, particularly serious VR games, was entirely beneficial for students with special needs, it should be noted that as each nascent educational or assistive aid develops, there is a dark side that should be considered along with its benefits. This is something that should be kept in mind whenever discussing the pros and cons of new technology. According to Bailenson (2018), the likelihood of having unfavorable mental and physical impacts increases when the use of fully immersive VR platforms is taken into consideration. The most typical symptoms of VR sickness include nausea, extreme tiredness, disorientation, and visual flashbacks. Certain tactics for authoring VR, like as micro-learning or learning chunks, may significantly reduce the severity of these symptoms. According to Kapp and Defelice (2019), the term "VR microlearning" describes the process of producing content in VR in distinct chunks that have predetermined durations and intervals. Because some materials do not have the capability to be authored in VR, this strategy is based on two guiding principles: first, the content that will be authored in VR should be easily distinguishable from other content, and second, the topic of the content should be such that it can be reflected accurately in an environment that supports fully immersive learning.

This present research is important for a number of reasons, including the following: (A) bridging the gap between distinct areas of study, such as TEFL, therapeutic education, and speech-language therapy through the use of novel objectives; (B) recruiting an understudied community of EFL students from Iran; (C) employing a quasi-Solomon four-group design to eliminate the possibility of adverse effects from pre-test sensitization; (D) assisting an understudied community of EFL students who (F) taking into account that the microlearning or learning nugget approach to using virtual reality technology is the most secure method for generating VR content, (G) providing interdisciplinary implications for various fields of science such as TEFL, assistive and educational technology, therapeutic education, and speech-language therapy. (F) considering measuring speaking skills through oral CAF which is a valid and reliable measurement scale for speaking skills, particularly among the ones with speech disorders. (G) considering measuring speaking skills through oral CAF which is a valid and reliable measurement scale for speaking skills, particularly among the ones with speech disorders.

Literature Review

The potential associations between the variables in this study are supported by four distinct theories: the Assistive Learning Theory (ALT), the Cognitive Theory of Multimedia Learning (CTML), the Speech and Cognitive Theory (SCT), and the Cognitive Load Theory (CLT). These theories help justify the rationals that lie behind the objectives that are the primary focus of this investigation. Using various current technology advances, such as VR learning solutions, may considerably enhance a variety of areas of learning difficulties, including verbal and nonverbal characteristics, as well as one-dimensional and multi-dimensional aspects. This is according to

ALT, which was published in 2021 by Green. The CTML claims that activating numerous sensory channels during learning enhances the possibility of acquiring the subject in focus, which is what occurs in IL, as visual, auditory, and haptic senses are engaged simultaneously, and the extra cognitive burden is avoided (Schunk, 2012). According to SCT (Thomas, 2021), there is an inverse correlation between the quantity of learning resources used in the brain during cognitive processes or functions and the degrees of communication abilities in terms of the complexity and fluency of their utilization. Therefore, effective virtual reality solutions assist users by reducing the amount of cognitive load that is required throughout the learning process and by providing the groundwork for correct and fluent communication. In addition, according to the CLT, efficient learning occurs when the quantity of extraneous cognitive load that must be managed is reduced. This results in an increase in the levels of intrinsic cognitive burden. In addition, Frehlich (2020) found that the amount of unnecessary cognitive burden in virtual reality (VR) learning settings was negatively correlated with the degree of immersion in the environment. With the help of Virtual Reality Head Mounted Displays (VR-HMDs), the participants in this investigation were given the opportunity to participate in a VR learning environment that was completely immersive.

Related Research

Vercellotti (2012) conducted an international longitudinal study to investigate whether or not participation in an EFL course focusing on oral CAF over the course of three months is beneficial to EFL learners. This study is considered one of the pioneer studies in the literature investigating the impacts of a 2D virtual reality environment on oral Complexity, Accuracy, and Fluency (CAF). During the intervention phase, which lasted for a total of three months, students of English as a foreign language were forced to utilize a 2D virtual-reality platform. This gave them the opportunity to view brief video clips on a variety of themes linked with the language learning process, including articulation system mechanics and difficult English sounds. After that, they were forced to do a series of group activities with the buddies they had brought with them. The results suggested that this setting had substantial impacts on all of the oral CAF measures, with the exception of one quality in accuracy (lexical diversity), and one quality in fluency (length of fluent run). Regarding the use of a virtual reality-based 2D game to boost the oral output of English native speakers, Zariski (2014) employed World of Warcraft, which is a well-known massively multiplayer online game (MMOG), in order to improve the oral complexity, accuracyand , and fluency of English native speakers. The participants were split up into four groups, each of which had an English native speaker and an EFL student. They were required to play World of Warcraft for five to six hours each day for a period of four weeks. The findings revealed that two EFL students had substantial gains in all of the oral CAF characteristics, whereas the other two students of EFL only saw an improvement in one or two of the oral CAF qualities. As a result of the interviews, it was clear that each and every one of the EFL students had a favorable opinion of this strategy, which was borne out by their participation rates.

In one of the few studies that are remotely linked to one another that have been carried out in the Iranian EFL setting, Bava Harji and Gheitanchian (2017) used a quasi-experimental research design to explore the effect of a 2D virtual-reality multimedia environment on EFL learners' oral CAF. The findings showed that these tactics were successful on the oral CAF among the Iranian EFL learners; nonetheless, a significant difference was discovered between these groups, as the people in the experimental group had superior performance regarding the CAF domains in comparison to their counterparts in the control group. Spring (2020) investigated the efficacy of short video production in a primary 2D virtual-reality environment to promote oral CAF among Chinese EFL learners. This was done with the intention of employing a collaborative activity in a

primary virtual-reality platform to potentially increase EFL learners' oral CAF. Certain domains to test each feature in the oral CAF paradigm were included in the interviews that were used in both the pre-intervention and post-intervention periods. The findings indicated that oral CAF measures improved after the intervention, although the difference between their performance in the pre-intervention and post-intervention phases in terms of oral CAF measures was not statistically significant. It has been claimed that more investigations are required to study the reasons why the solution that was deployed was not considerably beneficial. This is despite the fact that the efficacy of this virtual reality-based solution was validated to certain restricted levels. Therefore, it is possible that future study may concentrate on this tactic in a variety of English learning environments with a wide range of English learners.

Hashemifardnia et al. (2021) evaluated the use of a Massive Open Online Course (MOOC) on the oral CAF of Iranian EFL learners and their attitudes on the use of this platform to develop speaking abilities. This research was done in consideration of the usage of a 2D virtual-reality-based platform in the Iranian EFL setting. The intervention required its participants to view a series of video segments about speaking skills in concentration, complete a task, either by themselves or with the assistance of their classmates, and get feedback from their teacher over the course of an academic semester. The findings provide insight into the efficiency of this approach on oral CAF among Iranian EFL students. Aside from that, the participants' mindset toward making use of this platform was upbeat and enthusiastic. In the wake of other studies conducted in the Iranian EFL context, Nasri and Sepehri (2022) investigated the impact of a 2D virtual-reality web-based environment on Iranian EFL learners' oral CAF and their apprehension towards the use of a foreign language. The intervention phase involved the presentation of certain chapters from an English textbook known as "Speaking English Like an American" in a web-based learning environment. Within this environment, participants were able to engage in real-time conversation with one another, get feedback from their classmates or teacher, send voice messages, and even do internet searches. The people in the control group were instructed face to face and were given the identical course content as those in the experimental group in the typical learning setting. According to the results of the post-test, the individuals in the experimental group performed much better than those in the control group when it came to oral CAF. In addition, individuals who took part in the study who were assigned to the experimental group reported feeling less anxious about speaking a foreign language than those who were assigned to the control group.

The current study was conducted in order to find an answer to the following question, which was prompted by the fact that there was a gap in the previously reviewed studies and the fact that those studies had both positive and negative aspects:

> ***RQ.*** *Is there a significant difference between the effects of IL and Communicative Language Teaching on the oral CAF of Iranian pre-intermediate EFL learners with Expressive Language Disorder?*

Methodology

Design and Context

The quasi-Solomon four-group design (Tashakkori et al., 2020; Van Engelenburg, 1999) was selected for this study for a variety of reasons, including the goals of the study, the possible effects of pre-test sensitization, and the characteristics of the participants. Within this design, there are two experimental groups (G1 and G3) and two control groups (G2 and G4). Pre-test and post-test procedures are carried out for Group 1 and Group 2 in this design; however, only Group 1 is given the intervention. In addition, only post-tests are conducted for Groups 3 and 4, and only Group 3 is subject to any kind of intervention. In order to assess the real success of the

intervention, a comparison is made between Group 1 and Group 3, and then between Group 2 and Group 4. A representation of the quasi-Solomon gour-group design may be seen in Figure 1.

Figure 1

Quasi-Solomon Four-group Design

Group	Baseline assessment (Pretest score)	Intervention	Endline assessment (Posttest score)
G1	O1	X	O2
G2	O3		O4
G3		X	O5
G4			O6

A private speech-therapy consulting center in Tehran was selected as the context of this study. This consultation center was established to help individuals, primarily students dealing with their language disorder, and carry out a series of associated research. Most of the users of this center suffered from ELD and stuttering, followed by some cases of Dysarthria. As some benefactors funded this center, the required devices to conduct this study were available to a sufficient level.

Participants

The participants of this study were users of the private speech-therapy consulting center in which, at the same time, they enrolled in foreign language learning courses in different languages, such as German, English, and Turkish, which their speech therapist prescribed to trigger their motivation toward learning a new language and enhance their communication skills. Due to the research objectives, a purposive sampling procedure was followed in selecting pre-intermediate EFL learners who successfully passed the pre-intermediate level of the Oxford English File (Fourth Edition). Although their success in passing the pre-intermediate course was an indicator of their English proficiency, an Oxford Quick Placement Test (OQPT) was also administered to ensure their proficiency levels in English. Their ELD levers were confirmed by a panel of speech-language pathologists to be between mild to moderate levels, according to Alberta University Special Education Coding-Criteria (SECC), 2020/21. Finally, 36 Iranian pre-intermediate EFL learners with mild to moderate levels of ELD were purposively selected. It should be noted that attrition bias was controlled by classifying the dropped-out participant into voluntary and involuntary groups. The participants in the experimental groups experienced IL, and the ones in the control groups were taught the material based on the conventional CLT. Specific codes of conduct in research in social science (Weinbaum et al. 2018) were meticulously followed during the selection of the participants and carrying out this study.

Material

The content that was used in this investigation needed to satisfy three criteria: first, it needed to correspond to the participants' existing degrees of English proficiency; second, it needed to be segmentable into a variety of learning chunks; and third, it needed to be advantageous in terms of the participants' existing levels of oral CAF. The teaching material was, in the end, decided to be *Collins English for Life: Skills - Speaking*. This book is organized into twenty different modules, each of which is centered on one of five overarching concepts. The VR-authoring technique was based on the microlearning or learning chunks approach as the most secure VR-authoring option in terms of avoiding the experience of unwanted physical and mental consequences by the

participants. This was accomplished by breaking up the overall learning experience into smaller, more manageable chunks. In order to produce learning chunks out of the content that was the center of attention, a specific framework (Kohnke, 2023) was followed. Four learning chunks were constructed for each individual unit. The first of these learning chunks was concentrated on acquiring the communicative skill that was the primary topic of the unit, and the other three were speech-drills exercises that the participants were required to complete.

Instruments

Virtual Reality Head Mounted Displays (VR-HMD)

The first instrument was VR-HMD, a high-quality wearable display used by the participants in the experimental groups to experience IL with the assistance of two joysticks. HTC Vive Focus 3 was used in this study.

Oxford Quick Placement Test (OQPT)

The OQPT consists of 60 items constructed to examine participants' general English proficiency, mainly classifying them into proficiency levels ranging from beginners (0-17) to very advanced (54-60). In this study, only pre-intermediate EFL learners were selected. According to Hill and Parry (2014), individuals scoring between 30-39 are considered pre-intermediate EFL learners.

Oral Complexity Accuracy and Fluency Test

When it comes to the third instrument, the technique for determining the oral CAF level in Second Language Acquisition (SLA) is rather complicated (Norris & Ortega, 2009). In order to minimize any potential duplication in the measures taken for this research, the most valid and reliable test, developed by Housen et al. (2012), was employed. This exam consisted of six open-ended questions. In a well-known test from the past (Larsen-Freeman, 2009), the T-unit, which is the grammatical phrase that may be as short as feasible without losing its meaning, or the Type/Token Ratio (TTR) was used to quantify the oral complexity of the subject. Both syntactic complexity and syntactic variety are examined in more recent versions of the oral CAF test (i.e., Housen et al., 2012). These are the two broad categories that are measured. The ratio of the number of sentences to the number of Analysis of Speech Units (AS-Units) in an individual's oral output is one way to quantify the syntactic complexity of their speech. A single sentence that contains an independent clause together with its associated subordinate clauses is referred to as an AS-unit. The production of a total number of innovative phrases, clauses, and sentences across a range of tenses is the measure of a language's syntactic diversity. Two characteristics are taken into consideration while making an assessment of the precision. The first one is error-free sentences, which are expressed as a percentage value of the orally-produced phrases without mistakes in their choice, order, morphology, and syntax. This may be broken down further into four subcategories. The second factor is proper verb forms, which refers to the proportion of verbs that are appropriate in terms of the agreement, modality, and tense of their use. The level of fluency is evaluated using the same two standards as the other two criteria, which are referred to as Rate A and Rate B. Rate A is the number of syllables that are generated verbally in one minute, while Rate B is the same as Rate A except that it does not include any syllables, words, or sentences that are repeated or reformed. This test was validated by utilizing the words and sentences provided in the intervention phase, as well as by asking a panel of experts from the speech-language pathology and TEFL disciplines for their input. This approach was quite similar to the validity and reliability procedure that was utilized for the prior instrument. Concerning the consistency of the ratings given by the different raters, two individuals were chosen at random throughout the pre-intervention and post-intervention phases in order to determine whether or not there was a correlation between the ratings given and the actual performance of the participants.

Data Collection Procedure

In the first stage of collecting the required data, the necessary permissions were obtained from the heads of different departments, including speech consultation and language departments. Second, the parents of the participants under 18 were informed about their child's participation in this study. Third, the participants' levels of general English proficiency were re-assured by implementing the OQPT. After that, an SLP panel confirmed their ELD conditions to be mild to moderate based on Alberta University Special Education Coding-Criteria (SECC) 2020/21. Finally, 36 Iranian pre-intermediate EFL learners with mild to moderate levels of ElD were selected and were divided into four groups, including two experimental groups (G1 and G3) and two control groups (G2 and G4). During the intervention phase, the participants in the experimental group received treatment that was experiencing IL. In the control groups, they went under a series of conventional CLT sessions similar in number to the ones carried out for the experimental groups. The participants in the experimental group (G1 and G3) could use VR-HMDs to experience IL, and those in the control group had access to a DVD player used for listening to the conversations or speech drills. There were twenty 70-minute sessions in the intervention phase in which each session was based on teaching one unit of the material across four learning chunks; in those, one was distinguished for teaching the main lesson focusing on articulation or speaking skills, and three learning chunks were allocated for practicing speech drills. The immersion level in the IL was fully-immersive, and participants were placed in different virtual contexts themed as a museum, café, street corner, campus, and library. After experiencing the first learning nugget, which mainly introduced the communicative skill in focus followed by a conversation, three learning chunks were presented in speech drills to practice and perform what they have learned. The participants' oral CAF was measured before and after the intervention phase. The oral CAF test was used as the pre-test and post-test. One therapist or an SLP panel was present as an observer or assistant in selecting the participants, implementing the intervention, and examining oral CAF.

Data Analysis Procedure

Systematic Analysis of Language Transcripts (SALT) and Sampling Utterances and Grammatical Analysis Revised (SUGAR) software were used to assist researchers in the process of measuring oral CAF. Besides, the R statistical computing software was used to analyze the obtained data. In line with the specific design used in this study, certain guidelines (Van Engelenburg, 1999) were followed. At first, descriptive statistics were calculated to examine the score of the participants regarding oral CAF in the pre-test. After that, skewness and kurtosis values, the Shapiro-Wilk test, and the Kolmogorov-Smirnov test were calculated to evaluate the normality assumption and decide whether a parametric or non-parametric test should be used. Further, the homogeneity of variance was checked before conducting a one-way ANOVA to check if there was a significant difference between the participants in the pre-test. After that, descriptive statistics were used to have a quick outlook of the oral CAF scores in the pre-test and post-test. Next, the normality assumption was checked for the scores obtained in the post-test. Finally, an unpaired t-test and a paired t-test were conducted to check the possible influences of the pre-test sensitization and the actual effectiveness of the intervention.

Results

In spite of the fact that the participants had already been successful in completing a pre-intermediate level English language learning course, an OQPT was used to double-check their level of English proficiency. Their pre-intermediate level of general English competence was substantiated by the findings of the descriptive statistics (N = 36, M = 36.34, SD = 1.86). In

addition, both the lowest possible score, which was 31, and the highest possible score, which was 37, were within the pre-intermediate English competency range of 30–39 points (Hill & Parry, 2014). In addition, there was a range of 6 scores, and the variation was equal to 3 times the standard deviation. The design of this study (Tashakkori et al., 2020; Van Engelenburg, 1999) called for a specific statistical method to be carried out in order to answer the third research question, which aimed to determine whether or not there was a significant difference in the participants' oral CAF between those who participated in the IL and those who underwent a series of conventional CLT sessions. This comparison was made in order to answer the question. To begin, it was necessary to investigate the subjects' oral CAF as part of the pre-test. Because of the way that this research was designed, the pre-test was only given to two groups: G1, which was one of the experimental groups, and G2, which was one of the control groups. Neither of these groups received any treatment other than the pre-test. The normality assumption of the acquired data was tested before the participants' oral CAF was evaluated during the pre-test. This was done before the participants' oral CAF was evaluated. In order to study the mean scores as well as the values of skewness and kurtosis, descriptive statistics were computed and analyzed. It is important to highlight that, in accordance with the criteria that were chosen for this investigation, oral complexity was defined as syntactic diversity and complexity, oral accuracy was defined as error-free phrases and accurate verb forms, and oral fluency was defined as Rate A and Rate B. The relevant findings are shown in Table 1.

Table 1

Descriptive Statistics for Scores of Oral CAF in the Pre-test

	G*	N	M	SD	Skewness	Kurtosis
Syntactic Complexity	G1	9	1.752	.491	1.301	.346
	G2	9	1.742	.341	1.450	1.664
Syntactic Variety	G1	9	4.111	1.054	.552	-.546
	G2	9	4.444	1.130	.176	-1.171
Error-free Clauses**	G1	9	7.333	1.000	.107	-.643
	G2	9	7.111	1.269	.683	-1.251
Correct Verb Forms**	G1	9	14.555	1.013	.270	-.763
	G2	9	14.888	.781	.216	-1.041
Rate A	G1	9	85.777	2.108	.544	.159
	G2	9	84.666	2.000	-.244	-.844
Rate B	G1	9	35.444	2.068	-.553	-1.041
	G2	9	35.222	1.641	.419	-.832

Note. * G stands for the group. ** Percentages are converted into frequencies.

As seen in Table 1, the results show that the values of skewness and kurtosis for oral complexity, including syntactic complexity (G1=1.301, .346; G2=1.450, 1.644) and syntactic variety (G1=.552, -.546; G2=.176, -1.171) were in the range of -2 and +2 confirming that the gathered results followed the assumption of normality (Llaudet & Imai, 2022). Considering the results of oral accuracy, the values of skewness and kurtosis across error-free clauses (G1=.107, -.643; G2=.683, -1.251) and correct verb forms (G1=.270, -.763; G2= .216, -.1.041) posited between -2 and +2 shedding light on the fact that the obtained results did not violate the normality assumption. Finally, with a focus on oral fluency, the values of skewness and kurtosis for rate A (G1: .544; .159; G2: -.244, -.844) and rate B (G1: -.553, -1.041; G2: .419, -.832) placed in the range -2 and +2 indicating that the assumption of normality was followed. As the skewness and kurtosis values for all the groups across different constructs of oral CAF did not violate the assumption of normality, Kolmogorov-Smirnov and Shapiro-Wilk tests were utilized

in an attempt to guarantee that the results did not violate the normality assumption. Table 2 displays the associated results.

Table 2

Kolmogorov-Smirnov and Shapiro-Wilk Tests for Oral-CAF Scores in the Pre-test

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	G	Statistic	df	Sig.	Statistic	df	Sig.
Syntactic Complexity	G1	.288	9	.069	..785	9	.063
	G2	.217	9	.200*	.877	9	.146
Syntactic Variety	G1	.209	9	.200*	.889	9	.194
	G2	.208	9	.200*	.899	9	.248
Error-free Clauses	G1	.192	9	.200*	.917	9	.364
	G2	.254	9	.098	.810	9	.216
Correct Verb Forms	G1	.264	9	.071	.892	9	.208
	G2	.223	9	.200*	.838	9	.065
Rate A	G1	.134	9	.200*	.982	9	.974
	G2	.186	9	.200*	.952	9	.709
Rate B	G1	.218	9	.200*	.921	9	.400
	G2	.216	9	.200*	.941	9	.588

Note. *. This is a lower bound of the true significance. a. Lilliefors Significance Correction

As demonstrated in Table 4.14, all the significance values in the Kolmogorov-Smirnov test for the oral complexity, including syntactic complexity across G1 (p =.069) and G2 (p=.200) and for the syntactic variety across G1 (p=.200), and G2 (p=.200) were more than the critical level of .05. In addition, the significance values of the Kolmogorov-Smirnov test for oral accuracy which is consisted of error-free clauses across G1 (p=.200), and G2 (p=.098), and for correct verb forms across G1 (p=.071), and G2 (p=.200), were higher than the critical value of.05. Considering the p-values of the Kolmogorov-Smirnov test for oral fluency containing rate A across G1 (p=.200), and G2 (p=.200), and rate B across G1 (p=.200), and G2 (p=.200) were above the significance level of 0.05. Considering the outcomes of the Shairo-Wilk test, the significance values were shown to be more than the significance level of .05 for oral complexity, including syntactic complexity across G1 (p=.063) and G2 (p=.146) and syntactic variety across G1 (p=.194), and G2 (p=.248).

The p-values of the Shapiro-Wilk test for oral accuracy, including error-free clauses across G1 (p=.364), G2 (p=.216), and for correct verb forms across G1(p=.208), G2 (p=.065) were higher than the cut-off value of .05. Regarding the significance values of the Shapiro-Wik test for oral fluency including rate A across G1 (p=.974), and G2 (p=.709), and rate B across G1 (p=.400), and G2 (p=.588) were more than the critical value of .05. All in all, the results of the normality tests indicated the assumption of normality for oral CAF scores in the pre-test for G1 and G2 followed the assumption of normality (Llaudet & Imai, 2022). Considering the obtained results, a parametric test had to be selected to check if there was any significant difference between G1 and G2 in terms of oral CAF in the pre-test. Before conducting an ANOVA test, it was required to check the assumption of homogeneity of variances. Table 3 is the representation of the related outcomes.

Table 3

Test of Homogeneity of Variances for Scores of Oral CAF in the Pre-test

		Levine Statistic	Df1	Df2	Sig.
Syntactic	Based on mean	1.039	1	16	.323

Complexity	Based o median	.322	1	16	.578
	Based on the median with adjusted df	.322	1	14.649	.579
	Based on trimmed mean	.861	1	16	.367
Syntactic Variety	Based on mean	.211	1	16	.653
	Based o median	.105	1	16	.750
	Based on the median with adjusted df	.105	1	15.611	.750
	Based on trimmed mean	.249	1	16	.625
Error-free Clauses	Based on mean	.000	1	16	1.000
	Based o median	.000	1	16	1.000
	Based on the median with adjusted df	.000	1	16.000	1.000
	Based on trimmed mean	.000	1	16	1.000
Correct Verb Forms	Based on mean	1.217	1	16	.286
	Based o median	.457	1	16	.509
	Based on the median with adjusted df	.457	1	13.517	.510
	Based on trimmed mean	1.201	1	16	.289
Rate A	Based on mean	.008	1	16	.929
	Based o median	.000	1	16	1.000
	Based on the median with adjusted df	.000	1	16.000	1.000
	Based on trimmed mean	.008	1	16	.930
Rate B	Based on mean	.803	1	16	.383
	Based o median	.444	1	16	.514
	Based on the median with adjusted df	.444	1	14.400	.516
	Based on trimmed mean	.777	1	16	.391

As shown in Table 3, all the significant values were above the critical level (*0.05*); thus, the assumption of homogeneity of variances was not violated. After checking the required assumptions of normality and homogeneity of variances, a one-way ANOVA test was conducted to examine if there was any significant difference regarding the scores of the oral CAF in the pre-test. Table 4 indicates the associated results for G1 and G2.

Table 4

One-way ANOVA Test for Oral CAF Scores in the Pre-test

		Sum of Squares	**df.**	**Mean Square**	**F**	**Sig.**
Syntactic Complexity	Between Groups	.000	1	.000	.003	.961
	Within groups	2.867	16	.179		
	Total	2.867	17			
Syntactic Variety	Between Groups	.500	1	.500	.419	.527
	Within groups	19.111	16	1.194		
	Total	19.611	17			
Error-free Clauses	Between Groups	.000	1	.000	..000	1.000
	Within groups	16.000	16	1.000		
	Total	16.000	17			
Correct Verb Forms	Between Groups	.500	1	.500	.610	.446
	Within groups	13.111	16	.819		
	Total	13.611	17			
Rate A	Between Groups	5.556	1	5.556	1.316	.268
	Within groups	67.556	16	4.222		
	Total	73.111	17			
Rate B	Between Groups	.222	1	.222	.064	.804

Within groups	55.778	16	3.486
Total	56.000	17	

As presented in Table 4, the significant values across oral complexity, namely syntactic complexity (*p=.961*) and syntactic variety (*p=.527*); oral accuracy including error-free clauses (*p=1*), and correct verb forms (*p=.446*); and oral fluency consisted of rate A (*p=.268*), and rate B (*p=.804*) were higher than the significance level of 0.05, confirming that there were no significant differences between the G1 and G2 groups regarding oral CAF scores in the pre-test. In following the results of the oral CAF in the pre-test and post-test across all groups are indicated. Table 5 shows the related descriptive statistics.

Table 5

Descriptive Statistics for Oral CAF Scores in Pre-test and Post-test

	Pre-test				**Post-test**							
	G1		G2		G1		G2		G3		G4	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
SC*	1.75	.49	1.74	.34	3.80	.86	1.71	.41	3.62	.83	1.66	.23
SV*	4.11	1.05	4.44	1.13	7.77	.97	4.33	.86	7.77	.83	3.88	.78
EC**	7.33	1.00	7.11	1.26	15.44	1.13	6.88	1.05	15.11	.92	5.66	1.22
CV**	14.55	1.01	14.88	.78	23.33	1.22	14.22	.83	23.11	1.16	14.55	.72
Rate A	85.77	2.10	84.66	2.00	143.44	12.21	83.77	1.09	142.22	8.70	84.00	1.80
Rate B	35.44	2.06	35.22	1.64	78.44	6.42	35.77	1.64	78.88	7.09	35.88	1.53

Note. * SC, SV: Syntactic Complexity, Variety **EC: Error Free Cluases, CV: Correct Verb Forms

As shown in Table 5, the participants in one of the experimental groups (G1), which had passed the pre-test, showed better performance in oral CAF after the intervention phase, including syntactic complexity (*M=3.80, SD=.86*), syntactic variety (*M=7.77, SD=.97*), error-free clauses (*M=15.44, SD=1.13*), correct verb forms (*M=23.33, SD=1.22*), rate A (*M=143.44, SD=12.21*), and rate B (*M=78.44, SD=6.42*). On the other hand, the participants' performance in one of the control groups (G2) that experienced the pre-test did not reveal considerable improvements in the post-test across oral CAF, namely syntactic complexity (*M=1.71, SD=.41*), syntactic variety (*M=4.33, SD=.86*), error-free clauses (*M=6.88, SD=1.05*), correct verb forms (*M=14.22, SD=.83*), rate A (*M=83.77, SD=1.09*), and rate B (*M=35.77, SD=1.64*). So, based on the statistical procedures required by the design of this study to answer research questions, check the effects of pre-test sensitization, and examine the actual effectiveness of the intervention, a series of unpaired t-tests were computed regarding the scores of the oral CAF among the participants in all groups for the post-tests. To check if using the parametric test was an appropriate choice and to ensure the results of the post-tests followed the normality assumption, skewness and kurtosis values were computed following the calculation of Kolmogorov-Smirnov and Shapiro-Wilk tests. Table 6 displays the associated outcomes.

Table 6

Skewness, Kurtosis, and Tests of Normality for Oral CAF Scores in the Post-test

		Skewness	**Kurtosis**	**Kolmogorov-Smirnov[a]**			**Shapiro-Wilk**		
	G	Statistic	Statistic	Statistic	df	Sig.	Statistic	df	Sig.
Syntactic Complexity	G1	-.295	-.318	.129	9	.200*	.968	9	.882
	G2	1.690	1.262	.228	9	.198	.835	9	.061
	G3	.210	-.777	.137	9	.200*	.960	9	.803

	G4	-.341	-.826	.163	9	.200*	.924	9	.422
Syntactic Variety	G1	-.502	-.009	.257	9	.088	.903	9	.273
	G2	.660	.825	.317	9	.060	.973	9	.132
	G3	.501	-1.275	.269	9	.069	.808	9	.075
	G4	.216	-1.041	.223	9	.200*	.838	9	.065
Error-free Cluases	G1	.176	-1.171	.208	9	.200*	.899	9	.248
	G2	1.094	.611	.245	9	.127	.825	9	.069
	G3	.944	1.354	.325	9	.087	.846	9	.068
	G4	-.233	-1.556	.195	9	.200*	.870	9	.122
Correct Verb Forms	G1	.233	-1.556	.195	9	.200*	.870	9	.122
	G2	-.501	-1.275	.269	9	.069	.808	9	.075
	G3	.340	-1.579	.274	9	.070	.827	9	.091
	G4	1.014	.185	.333	9	.105	.763	9	.108
Rate A	G1	-.385	1.108	.167	9	.200*	.971	9	.905
	G2	.550	1.022	.308	9	.064	.866	9	.112
	G3	-.704	1.096	.181	9	.200*	.950	9	.695
	G4	-.165	.044	.178	9	.200*	.960	9	.799
Rate B	G1	-.272	-1.312	.140	9	.200*	.923	9	.414
	G2	.889	.348	.238	9	.151	.901	9	.259
	G3	-1.068	.289	.209	9	.200*	.856	9	.187
	G4	.235	-1.306	.163	9	.200*	.909	9	.308

Note. *. This is a lower bound of the true significance. a. Lilliefors Significance Correction

As seen in Table 6, all the values of skewness and kurtosis for expressive syntactic complexity G1 (*-.295, -.318*), G2 (*1.690, 1.262*), G3 (*.210, -.777*), and G4 (*-.341, -.826*) were in the range -2 and +2; in addition, all the values of Kolomogrov-Smirnove and Shapiro-Wilk tests across G1 (*.200, .882*), G2 (*.198, .061*), G3 (*.200, .803*), and G4 (*.200, .422*) were more than the critical level of 0.05 confirming that the assumption of normality was not violated regarding the post-test scores of syntactic complexity across four groups (Llaudet & Imai, 2022). Concerning the scores of syntactic variety across four groups, all the values of skewness and kurtosis for expressive vocabulary across G1 (*-.502, -.009*), G2 (*.660, .825*), G3 (*.501, -1.275*), and G4 (*.216, -1.041*) were in the range -2 and +2; besides, all the values of Kolomogrov-Smirnove and Shapiro-Wilk tests across G1 (*.088, .273)*, G2 (*.060, .132*), G3 (*.069, .075*), and G4 (*.200, .065*) were more than the critical level of 0.05 confirming that the assumption of normality was followed regarding the post-test scores of syntactic variety among four groups (Llaudet & Imai, 2022).

With a focus on the scores of error-free clauses among four groups, all the values of skewness and kurtosis across G1 (*.176, -1.171*), G2 (*1.094, .611*), G3 (*.944, 1.354*), and G4 (*-.233, -1.556*) were in the range -2 and +2; furthermore, all the values of Kolomogrov-Smirnove and Shapiro-Wilk tests across G1 (*.200, .248*), G2 (*.127, .069*), G3 (*.087, .068*), and G4 (*.200, .122*) were more than the critical level of 0.05 indicating that the normality assumption was followed regarding the scores of error-free clauses across four groups in the post-test (Llaudet & Imai, 2022). Regarding the scores of correct verb forms across four groups, all the values of skewness and kurtosis for expressive vocabulary across G1 (*.233, 1.556*), G2 (*.501, 1.275*), G3 (*.340, -1.579*), and G4 (*1.014, .185*) were in the range -2 and +2; besides, all the values of Kolomogrov-Smirnove and Shapiro-Wilk tests across G1 (*.200, .122*), G2 (*.069, .075*), G3 (*.070, .091*), and G4 (*.105, .108*) were more than the critical level of 0.05 confirming that the assumption of normality was not violated regarding the post-test scores of correct verb forms among four groups (Llaudet & Imai, 2022).

In regard to the scores of rate A among the four groups, all the values of skewness and kurtosis across G1 (*-.385, 1.108*), G2 (*.550, 1.022*), G3 (*-.704, 1.096*), and G4 (*-.165, .044*) were

in the range -2 and +2; furthermore, all the values of Kolomogrov-Smirnove and Shapiro-Wilk tests across G1 (*.200, .905*), G2 (*.064, .112*), G3 (*.200, .695*), and G4 (*.200, .799*) were more than the critical level of 0.05 indicating that the normality assumption was followed regarding the scores of rate A across four groups in the post-test (Llaudet & Imai, 2022). Regarding the scores of rate B across four groups, all the values of skewness and kurtosis for expressive vocabulary across G1 (*-.272, -1.312*), G2 (*.889, .348*), G3 (*-1.068, .289*), and G4 (*.235, -1.306*) were in the range -2 and +2; moreover, all the values of Kolomogrov-Smirnove and Shapiro-Wilk tests across G1 (*.200, .414*), G2 (*.151, .259*), G3 (*.200, .187*), and G4 (*.200, .308*) were more than the critical level of 0.05 confirming that the assumption of normality was not violated regarding the post-test scores of rate B among four groups (Llaudet & Imai, 2022). With the use of a parametric test, a series of unpaired sample t-tests were used to check the potential effects of the pre-test sensitization on the post-test scores across all groups and in terms of Oral CAF. Besides, in this regard, it was possible to check the actual effectiveness of the intervention on the construct and subconstructs in focus. Table 7 indicates the mentioned results.

Table 7

Unpaired T-test for Oral CAF Scores in Post-test across All Groups

		t	Sig.(2tailed)	M Difference	Std. Error Diff.	Lower	Upper
SC*	G1G2	6.575	<.001	2.095	.31871	1.419	2.771
	G1G3	.466	.648	.186	.400	-.662	1.036
	G1G4	7.195	<.001	2.141	.297	1.510	2.771
	G2G3	-6.129	<.001	-1.908	.311	-2.569	-1.248
	G2G4	.288	.777	.045	.158	-.289	.380
	G3G4	6.744	<.001	1.954	.289	1.340	2.568
SV*	G1G2	7.938	<.001	3.444	.433	2.524	4.364
	G1G3	.000	1.000	..000	.426	-.904	.904
	G1G4	9.354	<.001	3.888	.415	3.004	4.770
	G2G3	-8.598	<.001	-3.444	.400	-4.293	-2.595
	G2G4	1.143	.270	.444	.388	-.379	1.268
	G3G4	10.211	<.001	3.888	.380	3.081	4.696
EC**	G1G2	16.606	<.001	8.555	.515	7.463	9.647
	G1G3	.684	.504	.333	.487	-.700	1.366
	G1G4	17.600	<.001	9.777	.555	8.600	10.955
	G2G3	-17.564	<.001	-8.222	.468	-9.214	-7.229
	G2G4	2.269	.067	1.222	.538	.080	2.364
	G3G4	18.439	<.001	9.444	.512	8.358	10.530
CV**	G1G2	18.451	<.001	9.111	.493	8.064	10.169
	G1G3	.394	.699	.222	.563	-.973	1.417
	G1G4	18.493	<.001	8.777	.474	7.771	9.784
	G2G3	-18.600	<.001	-8.888	.477	-9.902	-7.875
	G2G4	-.905	.379	-.333	.368	-1.114	.447
	G3G4	18.675	<.001	8.555	.458	7.584	9.526
Rate A	G1G2	14.592	<.001	59.666	4.088	50.998	68.334
	G1G3	.244	.810	1.222	4.999	-9.376	11.821
	G1G4	14.440	<.001	59.4444	4.116	50.717	68.171
	G2G3	-19.996	<.001	-58.444	2.922	-64.640	-52.248
	G2G4	-.316	.756	-.222	.702	-1.711	1.267
	G3G4	19.658	<.001	58.222	2.961	51.943	64.500
R	G1G2	19.303	<.001	42.666	2.210	37.980	47.352
a							

G1G3	-.139	.891	-.444	3.190	-7.208	6.320
G1G4	19.326	<.001	42.555	2.201	37.887	47.223
G2G3	-17.756	<.001	-43.111	2.427	-48.258	-37.964
G2G4	-.148	.884	-.111	.749	-1.699	1.477
G3G4	17.766	<.001	43.000	2.420	37.869	48.130

Note. * SC, SV: Syntactic Complexity, Variety **EC: Error Free Cluases, CV: Correct Verb Forms

As seen in Table 7, in all components of oral CAF, only the interactions of G1-G3 ($p>0.05$), which were experimental groups, and G2-G4 ($p>0.05$) that were control groups showed no significant differences, but all the other groups were significantly different from each other in terms of all the components of oral CAF scores in the post-test ($p<0.05$). Thus, it can be concluded that the pre-test sensitization was successfully managed, and the intervention improved the participants' oral CAF in the experimental groups. It was required to use a paired sample t-test to compare the pre-test and post-test scores across G1 and G2 groups to check if the IL significantly enhanced oral CAF among the Iranian pre-intermediate EFL learners with ELD. Table 8 indicates the related outcomes.

Table 8

Paired T-test for the Oral CAF Scores across Pre-test and Post-test Between G1 and G2

			Paired Differences							
						95% Confidence Interval of the Difference				
			M	SD	SEM	Lower	Upper	t	df	Sig. (2tailed)
G1	Pre-test, Post-test	SC*	-2.05	1.09	.363	-2.89	-1.21	-5.65	8	<.001
		SV*	-3.66	1.73	.577	-4.99	-2.33	-6.35	8	<.001
		EC**	-8.11	1.36	.454	-9.15	-7.06	-17.8	8	<.001
		CV**	-8.77	1.71	.571	-10.09	-7.45	-15.3	8	<.001
		Rate A	-57.6	12.86	4.28	-67.55	-47.77	-13.4	8	<.001
		Rate B	-43.0	7.44	2.48	-48.72	-37.27	-17.3	8	<.001
G2	Pre-test, Post-test	SC*	.028	.392	.130	-.272	.330	.221	8	.831
		SV*	.111	.781	.260	-.489	.712	.426	8	.681
		EC**	.444	1.740	.580	-.893	1.781	.766	8	.466
		CV**	.666	1.00	.333	-.102	1.435	2.00	8	.081
		Rate A	.888	2.57	.857	-1.08	2.86	1.03	8	.330
		Rate B	-.555	2.74	.914	-2.66	1.55	-.607	8	.560

Note. * SC, SV: Syntactic Complexity, Variety **EC: Error Free Cluases, CV: Correct Verb Forms

As revealed in Table 8, regarding the scores of the oral CAF, including syntactic complexity ($p<.001$), syntactic variety ($p<.001$), error-free clauses ($p=.017$), correct verb forms ($p<.001$), rate A ($p<.001$), and rate B ($p<.001$), there was a significant difference among the participants in G1 ($p<.05$) who experienced IL between the pre-test and post-test. Besides, concerning the scores of the oral CAF, namely syntactic complexity ($p=.831$), syntactic variety ($p=.681$), error-free clauses ($p=.466$), correct verb forms ($p=.081$), rate A ($p=.330$), and rate B ($p=.560$), there was no significant difference among the participants in G2 ($p>.05$) who were taught based on the conventional CLT method between the pre-test and post-test. Thus, it can be concluded that IL significantly enhanced all constructs of the oral CAF among the Iranian pre-intermediate EFL learners with ELD.

Discussion

The above results demonstrated that there was a considerable difference between the impacts of IL and those of traditional education on the oral CAF of Iranian pre-intermediate EFL students who were learning English for speakers of other languages who had ELD. Oral CAF consists of three components: complexity, accuracy, and fluency. The first component comprises the measuring of syntactic complexity and variation; the second component involves the examination of error-free phrases and accurate verb forms; and the third component involves rate A and rate B indexes. The findings are consistent with those of other research (Vercellotti, 2012; Zariski, 2014), according to which the use of VR technology as a teaching platform helped international EFL learners in terms of oral CAF, while the findings are contradictory in terms of syntactic variety and rate A.

The findings of this study are in accordance with those of the other studies. Syntactic variety is an index that measures the diversity of the sentences, clauses, or phrases that are created depending on their tenses, and rate A is a measure of the number of syllables that are produced in a certain time period. Due to the fact that the participants in the aforementioned studies spoke different mother tongues, such as Arabic, Korean, Indian, and Chinese, the discrepancy that emerged between the findings of this study and those of Vercellotti (2012) may have its origins in something that is known as a trade-off effect among EFL/ESL learners who speak different mother tongues when learning English. This effect occurs when various aspects of the mother tongue, particularly the syntactical features, significantly influence the syntactic aspects of theIn contrast to the findings from Spring (2020), which found that the VR learning platform did not improve Chinese EFL learners' oral CAF based on an experimental study, these data show that the platform does make a difference. The disparate findings may be due to differences in the impacts of the Virtual Reality (VR) technology employed in the study conducted by Spring (2020) compared to the effects of the 3D VR-infused training used in this research on Verbal Working Memory (VWM) and the quantities of extraneous cognitive load. The study was conducted in 2020.

According to Gathercole and Baddeley (2014), verbal working memory (VWM) refers to a passive temporary ability to store verbal knowledge. This capacity is favorably associated with the functional power of working memory. Jones et al. (2022) found that employing completely immersive or semi-immersive 3D technology was more effective in boosting working memory capacity than utilizing non-immersive and desktop-based 2D virtual-learning platforms. In addition, fully immersive or semi-immersive 3D technology was more beneficial than non-immersive desktop-based 2D virtual-learning platforms. Therefore, it is possible to draw the conclusion that the use of 2D VR technology did not succeed in enhancing VWM as a component of working memory among Chinese EFL students, which resulted in no substantial improvement in their oral CAF. In addition, it has been shown in a few studies (Roettl and Terlutter, 2018; Yin et al., 2023) that users of 2D VR technology experience much greater levels of extraneous CL compared to users of 3D VR platforms. Because of this, it is possible that the contradiction between the findings of this research and those of Spring (2020) was caused by the ineffectiveness of 2D VR technology in enhancing VWM and by imposing high levels of extraneous CL on the user of these settings. This conclusion is reached as a consequence of the arguments that were given, which led to the conclusion.The findings are consistent with those of other research that were carried out in the Iranian EFL context (Bava Harji and Gheitanchian, 2017; Hashemifardnia et al., 2021; Nasri and Sepehri, 2022). In those studies, the VR learning platform considerably improved the oral CAF of the learners of Iranian EFL. According to the theory of embedded learning, a critical theory in the design of instructional materials, the potential reason accounting for the obtained outcomes regarding the significant effectiveness of

VR learning platform on Iranian EFL learners' oral CAF is justifiable. This theory is rooted in the theory of embedded cognition. According to the theory of embedded cognition, humans have access to a variety of resources for cognitive awareness in addition to their brain alone. These resources include the individual's body, their environment, and the connections or interactions that exist between these resources (Shapiro, 2019). According to the embedded learning theory, which was structured on the basis of the earlier central theory, effective learning takes place when the instructional material requires learners to use a variety of sensory channels. These sensory channels are primarily activated when learners are placed in authentic settings or collaborate with their peers to construct meaning out of information that has been presented (Molnar, 2020). This theory was recently put into practice in the field of educational virtual reality technology, confirming the outcome by highlighting the fact that virtual reality (VR) learning platforms, particularly immersive and semi-immersive ones, providing an authentic environment in which different sensory channels of users are engaged significantly increases the effectiveness of teaching material (Lan & Grant, 2021).

Conclusion

The findings of this research demonstrated that using an effective quasi-Solomon four-group design successfully removed pre-test sensitization and that employing IL greatly boosted Oral CAF among Iranian pre-intermediate EFL learners with ELD in comparison to standard CLT. Both of these findings were supported by the findings of the study. In addition, the findings of this research show that utilizing IL, especially immersive and completely immersive ones, may improve VR users' Verbal Working Memory (VWM) and lower the quantities of unnecessary CL. This is in relation to the increased oral CAF. This research has the potential to improve TEFL, therapeutic education, educational technology, and psychology by shining light on the great advantages of utilizing IL significantly for those with learning or verbal difficulties. Because this study is one of the pioneer studies in the literature, it may aid these fields. It is possible that future research will follow the same objectives, but with different types of learners who have different levels of English proficiency (intermediate or advanced), different types of verbal disorders (such as stuttering or dysarthria), and even the most basic brain-induced disorders (such as autism spectrum disorder or ASD), as well as different research designs (sequential mixed-method research).

References

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. WW Norton & Company.

Bava Harji, M., & Gheitanchian, M. (2017). Effects of multimedia task-based teaching and learning approach on EFL learners' accuracy, fluency and complexity of oral production. Turkish Online Journal of Educational Technology-TOJET, 16(2), 25-34.

Cockrell, H. (2021). *Learning disabilities & disorders definition, facts, its effects & more: learning disabilities and related disabilities*. Independent Publication.

Cook, A. M., & Polgar, J. M. (2014). *Assistive technologies: principles and practice*. Elsevier.

Frehlich, C. (2020). *Immersive learning: A practical guide to virtual reality's superpowers in education*. Rowman & Littlefield Publishers.

Gathercole, S. E., & Baddeley, A. D. (2014). *Working memory and language*. Psychology Press.

Green, J. L. (2021). *Assistive technology in special education: Resources to support literacy, communication, and learning differences*. Routledge.

Greengard, S. (2019). *Virtual reality*. MIT Press.

Hashemifardnia, A., Shafiee, S., Esfahani, F. R., & Sepehri, M. (2021). Effects of massive open online course (MOOC) on Iranian EFL learners' speaking complexity, accuracy, and

fluency. *Computer-Assisted Language Learning Electronic Journal (CALL-EJ), 22*(1), 56-79.

Hill, C., & Parry, K. (2014). *From testing to assessment: English an international language.* Cambridge.

Housen, A., Kuiken, F., & Vedder, I. (2012). *Complexity, accuracy and fluency: Dimensions of L2 performance and proficiency.* John Benjamins.

Kaplan-Rakowski, R., & Gruber, A. (2022). The impact of high-immersion virtual reality on foreign language anxiety when speaking in public. *TechTrends, 24*(6), 56-69. *https://doi.org/10.1007/s11528-022-00717-w*

Kapp, K. M., & Defelice, R. A. (2019). *Microlearning: Short and sweet*. American Society for Training and Development.

Kecskes, I., & Nde Papp, T. (2000). *Foreign language and mother tongue*. Psychology Press.

Lan, Y. J., & Grant, S. (2021). *Contextual language learning: real language learning on the continuum from virtuality to reality*. Springer Nature.

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, *30*(4), 579-589. https://doi.org/10.1093/applin/amp043

Maniam, M., & Kesevan, H. V. (2016). Mother tongue syntactical influence in a ESL class. *International Journal of Language and Linguistics*, *3*(3), 46-50.

Mayer, R. E. (2009). *Multimedia Learning.* Cambridge.

Mohammad, T. (2015). Phonological, morphological and syntactical influences of mother tongue on Arab learners. *English Review: Journal of English Education, 4*(1), 54-65.

Molnar, C. (2020). *Interpretable machine learning*. Lulu Press.

Nasri, M., & Sepehri, M. (2022). An account of Iranian EFL learners' speaking complexity, accuracy, and fluency and foreign language speaking anxiety in a computer-assisted language learning environment. *Computer-Assisted Language Learning, 23*(2), 299-321.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, *30*(4), 555-578.

Roettl, J., & Terlutter, R. (2018). The same video game in 2D, 3D or virtual reality–How does technology impact game evaluation and brand placements?. *Plos One, 13*(7), 200-217. https://doi.org/10.13712Fjournal.pone.0200724

Sabry, F. (2022). *Virtual reality: Taking the concept of augmented reality to the next level by creating a fully computer-generated simulation of a different world.* One Billion Knowledge.

Schunk, D. H. (2012). *Learning theories an educational perspective*. Pearson Education.

Shapiro, L. (2019). *Embodied cognition*. Routledge.

Spring, R. (2020). Can video-creation project work affect students' oral proficiency? An analysis of fluency, complexity and accuracy. *TESL-EJ, 24*(2), 25-36.

Tashakkori, A., Johnson, R. B., & Teddlie, C. (2020). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage Publications.

Thomas, L. (2021). *Manual of cognitive treatment for adults: Simple speech solutions*. Newman Springs.

Van Engelenburg, B. (1999). Statistical analysis for the solomon four-group design. *OMD Research, 14(*2), 1-37.

Vercellotti, M. L. (2012). *Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other*. Pittsburgh Publication.

Weinbaum, C., Landree, E., Blumenthal, M. S., Piquado, T., & Gutierrez, C. I. (2018). *Ethics in scientific research: An examination of ethical principles and emerging topics.* Routledge.

York, J. (2019). *Language learning in complex virtual worlds: Effects of modality and task complexity on oral performance between the virtual world and face-to-face tasks*. University Press.

Westwood, P. (2017). *Learning disorders: A response-to-intervention perspective.* Taylor and Francis.

Yin, Y., Sun, J., Liu, Y., Wang, H., & Jing, P. (2023). The cognitive load of observation tasks in 3D video is lower than that in 2D video. *Arxive, 19*(3), *302-319.* https://doi.org/10.48550/arXiv.2302.12968

Yudintseva, A. (2023). Virtual reality affordances for oral communication in English as a second language classroom: A literature review. *Computers & Education: X Reality*, *2(9)*, 100-118.

Zariski, R. (2014). *World of Warcraft as a medium for intermediate-level English language acquisition: Leveling up accuracy, fluency, and lexical complexity*. Victoria Publication.

Biodata

Ms. Mojgan Khaki is a Ph.D. candidate of TEFL and lecturer at the Department of English, Isfahan (Khorasgan Branch), Islamic Azad University, Isfahan, Iran. Her research interests include Teaching Translation, Critical Discourse Analysis, and English Literature.
Email: *Mojgan.khaki96@gmail.com*

Dr. Hossein Heidari Tabrizi is an associate professor of TEFL and currently the head of the English Department at Islamic Azad University, Isfahan Branch, Isfahan, Iran. He is the director-in-charge of Research in English Language Pedagogy (RELP). His research interests include language assessment, translation studies, and critical discourse analysis.
Email: *heidaritabrizi @gmail.com*

Dr. Sajad Shafiee is an Assistant Professor of Applied Linguistics at the Department of English, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran. He has published a good number of articles on testing and research issues in local and international journals. His research interests include testing, research, and materials development.
Email: *s.shafiee@iaushk.ac.ir*