

The Effect of Drama-based Method on Pre-intermediate EFL Learners' Speaking Skill and Elementary EFL Learners' Speaking Fluency, Comprehensibility, Grammaticality, and Willingness to Communicate

¹Behzad Barekat*

²Hosna Nazemi

IJEAP-2006-1569

Received: 2020-06-27

Accepted: 2020-09-14

Published: 2020-09-22

Abstract

Over years, there has always been a debate about how different methods and techniques improve speaking ability of foreign language learners. In this regard the present study examined two separate treatments. The first treatment was conducted for 20 sessions to examine the effects of "drama-based method of teaching" or more concisely "drama-based method" to enhance the speaking skill of 20 non-English major Iranian students; their English proficiency level based on the scores they gained in a standard proficiency test was pre-intermediate. Role-play was used for pre/posttests. To examine if a drama-based method improves elements of speaking skill _including fluency, comprehensibility, grammaticality, and willingness to communicate_ in learners with a lower level of language proficiency, a second treatment was conducted with 30 Iranian students whose major was not English and their English proficiency level was Elementary. In both treatments, participants took an oxford placement test to reveal if their English language proficiency was were at the intended levels. The same kind of role play test, which was introduced by Bartz, was used as the pretest/posttest. Data collection was quantitative and data analyses were conducted through paired t-tests in the first treatment and Mann-Whitney and Wilcoxon tests in the second treatment. The findings of the first treatment revealed improvement in speaking skill and the second treatment confirmed that each intended element of speaking skill (fluency, grammaticality, comprehensibility, and willingness to communicate) was enhanced. Being related to real-life situations, drama-based activities are more purposeful and motivate the learners more than usual.

Key words: Drama-based Method, Speaking Skill, Comprehensibility, Grammaticality, Willingness to Communicate

1. Introduction

There are many teaching methods and techniques used in language teaching to provide connections to daily life activities. Making use of the drama-based method is one of the effective recent methods in EFL (Bite, 2012; Chuhan, 2004; Comajoan, 20014). On this basis, this study aimed to provide a number of Iranian language learners with a variety of drama-based tasks in order to examine whether they can make a connection between language features and real-life communication. The activities are student-centered and give more time to students for communication in English. The chance for speaking is provided through dramatized activities including, improvisation, role plays, mime, and simulation. The topics are meaningful and related to student's needs and real-life communicative activities, speaking skill included

According to Hu (1996) and Chauhan (2004), effective communication is difficult for the learners in a situation they call "dump English", a term referring to the situation when students want or need to communicate in English but they can't perform the task successfully due to reasons such as tension, shyness, or lack of communicative competence in English (cited in Shen & Suwanthep, 2011, p. 2). In countries like Iran, English language is learned as a foreign language and the scarce

¹Associate Professor in Comparative Literature (Corresponding Author), behzadbarekat@yahoo.com; Department of English Language, Guilan University, Guilan, Iran.

²MA in TEFL, nezami.july24@gmail.com; Parand Language Institute, Guilan, Iran.

expose of learners to English can be the source of those already mentioned shortcomings for effective oral communication.

Real communication involves ideas, emotions, feelings, appropriateness and adaptability (Chauhan, 2004, p.1). The conventional language classes hardly give learners an opportunity to use language in this manner and also develop fluency. For these reasons, it is so important for English teachers to find effective techniques to enhance learners' speaking abilities. The present study, in its second treatment, tried to focus on and examine four parameters which play prominent role in speaking skill including; fluency, comprehensibility, grammaticality and willingness to communicate. Drama-based tasks provide situations and activities which let learners practice speaking involving most of the factors existing in real life communication without any risk and danger. Thus, they may lead to improvements in fluency, willingness to communicate, grammaticality and comprehensibility.

This paper focuses on using drama-based tasks to instruct EFL learners in 2 separate treatments. The first treatment conducted in 2018, tried to test if drama-based method improves pre-intermediate learners' speaking skill. One of the wrong beliefs toward using drama method in language learning environment is that learners must be proficient enough in target language in order to act in drama-based situations (Lai-wa Dora To, Yuk-lan Phoebe Chan, Yin Krissy Lam & Shuk-kuen Yvonne, 2011; Suganda, Zuraida, & Kurniawan, 2017). Consequently, second treatment conducted in 2019, tried to test if drama-based activities improve elementary learners' speaking skill. Moreover, in second treatment, four speaking elements_ fluency, comprehensibility, grammaticality and willingness to communicate_ were analyzed separately to find out whether drama method enhances each of intended elements independently. The research questions are as follows:

Research Question One: Does drama-based method significantly improve speaking skill of pre-intermediate EFL learners?

Research Question Two: Does drama-based method significantly improve each elements of speaking, i.e. fluency, comprehensibility, willingness to communicate, and grammaticality of elementary EFL learners?

2. Review of literature

2.1. Drama-based Method

Wessels (1987, p. 7) defined drama as a normal situation in daily life: "Drama is doing. Drama is being. Drama is such a normal thing. It is something that we engage in daily when we face with difficult situations." The word drama is originally from the Greek which means "to do, to act." Action is an essential part of learning a language, as it develops body language, increases motivation, and keeps students involved in the learning process (Gorjian, Moosavinia, & Jabripour, 2010; Miccoli, 2003). Drama also creates a friendly, stress-free atmosphere where optimal learning occurs (Conejeros & Fernandez, 2008; Gorjian, Moosavinia & Jabripour, 2010; Miccoli, 2003).

Wessels (1987) also emphasized on communication skills and believed a drama-based activity is one that causes genuine communication, a way of teaching that takes into account the background, emotions, the relations, the status, the body language and other paralinguistic features (cited in Comajoan, 2014, pp.9). Using drama, includes different activities which can be used in class, such as simulation, role play, mime, improvisation, reading plays and so on. All of these activities cause communicative language learning which is advocated by Morrow (1981) and Brumfit (1984) (cited in Vasantao, 2012, p.6). The following elements of speaking that will be investigated in this study:

2.1.1. Fluency

Hartmann & Stork (1976, cited in Brown, 1996, p.60) claimed that a person is a fluent speaker of a language when he is able to use structures accurately while he is concentrating on the content and uses the units automatically as in a normal conversation. Koponen and Riggensbach (2000, p.6)

define fluency as a performance-based phenomenon which is related to the flow, continuity, automaticity or smoothness of the speech. Fillmore (1979, p.93) stated four abilities that might be embodied under the term fluency, the first is the ability to talk at length with few pauses.

According to Gatbonton and Segalowitz (1988), Maurice (1983), Schneider (1993) the following principles need to be considered when designing and doing fluency building activities:

- To Incorporate repetition
- To increase the amount of speaking time
- To allow time to prepare before speaking
- Using familiar and motivating topics (cited in Kellem, 2009)

2.1.2. Grammaticality

Richards, Platt and Weber (1985) stated that 'grammar is a description of the structure of a language and the way in which linguistic units such as words and phrases are combined to produce sentences in the language' (Richards, Platt and Weber, 1985, p. 97). Crystal (1992) defines grammaticality as the 'conformity of a sentence or part of a sentence to the rules defined by a particular grammar of the 'language' (Nunan, 2001, p. 35).

Nunan (1999) claimed that while grammar lessons can work for some people, it is not as effective for others. However, opportunity to use the language can benefit all the learners. According to Aristotle, "the things we have to learn before we do them, we learn by doing them." The philosophy "learning by doing" (Dewey, 1938) embedded in CLT has provided a valuable insight for grammar teaching and it said formal instruction should incorporate activities of language use so that their learners can put theories into practice". (Wu, C. H., 2007, p.62)

2.1.3. Willingness to Communicate

Macintyre et al. (1998, p.547) defined willing to communicate in L2 as "a readiness to enter into discourse at a particular time with a specific person or persons, using a L2". So, WTC is the intention to initiate communication. Scholars suggest that conversational interaction is an essential part of learning a L2 so one of the fundamental goals of L2 instruction should be producing learners who are willing to use the language for authentic communication so producing WTC (which influences not only speaking mode but also listening, writing and reading modes) is a crucial component of modern language instruction, and therefore, the goal of the learning process should be to increase language students' willingness to communicate (Macintyre et al., 1998). Kang (2005) claimed that more interaction causes more language development and learning and situational willingness to communicate (WTC) in a second language (L2) emerge and fluctuate during a conversation situation (cited in Mahmoodi & Moazam, 2014, p.1069). According to Dougill (1987) and Teylor (2000), drama techniques can satisfy the needs of language learners in that they create motivation, enhance confidence, and provide context in learning a language (cited in Janudom & Wasanasomsithi, 2009, p. 4). Ladousse (1987, p.3) stated that role play is a means of increasing students' motivation, engagement, and confidence. According to Amato (1998, p.145) the use of drama into language teaching causes students lose themselves in the characters, plots and situations, and also leads to a decrease in anxiety levels and heightened students' self-confidence, self-esteem and awareness. McCroskey and Richmond (1987) conceptualized WTC as a trait but asserted that sometimes situational variables can impact on it. Reflecting on this trait view, researchers tried to find out those individual's variables which affect WTC. The results indicated that variables like [...] , and motivation influence on WTC (Donovan & MacIntyre, 2004; Hashimoto, 2007; MacIntyre, 1994).

2.1.4. Comprehensibility

Bartz (1979) defined comprehensibility as the ability of the student to make himself understood, to convey meaning. Practicing language in real life situations and interactions happen in drama method. So, learners should get the ability to interpret and respond to nonverbal clues (like facial expression and tones which happens in activities like pantomime, role play and play reading). That

is to say, language learners need to practice and have knowledge not only the linguistic forms but also other communication skills and devices to "reach mutual understanding". Through these interactions, learners figure out what they need to do to keep the conversation going and make themselves comprehensible (Lightbown and Spada, 2002, p.43). Additionally, Via (1987) defined drama techniques as strategies to communicate or convey the intended meaning and make one's comprehensible which includes a wide range of activities (Cited in Janudom & Wasanasomsithi, 2009, p.4).

3. Methodology

3.1. Participants

We conducted the intended treatment two times _once in the year 2016 and another time in 2018_, there were thus, two control groups and two experimental groups. In the first treatment, 2016, we aimed to check whether the general speaking ability of the learners will be significantly improved. There were 20 students at the pre intermediate level. Control group included 10 male students whose age ranged from 11 to 20, and they were learning English in Asr-e-Zaban institute in Astan-e-Ashrafie. Students in experimental group were 10 (8 females, 2 males) learners who were chosen through interview and oxford placement test from 40 volunteers that were requested to take part in free English language class and were informed through advertisement in Rasht (in the age range of 14-27). The institute in which experimental group was instructed was Boostan institute which was an English institution and located in Gilan, Rasht.

In the second treatment, 2018, total participants were 30 students at the elementary level. Control group included 15 (10 females, 5 males) students whose age ranged from 15 to 17. Students in experimental group were 15 female learners whose age ranged from 13 to 16. All participants were chosen through oxford placement (OPT) test from 150 students in Atlas language institute, Gilan, Rasht.

3.2. Instruments and Materials

The instruments which were used to fulfill the aims of the present study included:

- Advertising leaflet, Divar application, Instagram, and Facebook.
- Oxford Placement Test (OPT).
- A test in interview-role play format as pretest and posttest was used to stimulate participants to speak in real-life situations. Participants' speech was evaluated by considering factors such as, fluency, willingness to communicate, grammaticality, and comprehensibility. Bartz (1979)'s speaking test which is a communicative test was used in this study (Appendix A). To justify using such test, we bring some parts of his explanation here: "...Speaking in a foreign language involves more than being able to pronounce correctly, to put the stress on the correct syllable, or to supply the correct verb forms or endings. Rather, oral communication involves the combination of many skills, and depends on the student's ability to synthesize-not just to produce in isolation-discrete linguistic components... The development of the oral skills in foreign language learning can roughly be viewed as a process involving two basic levels-the linguistic level, in which the student acquires the skills necessary to perceive and produce grammatically correct utterances, and the communication level, in which the student acquires the skills necessary to carry out oral communicative interactions...in order for a test to be a true measure of communicative ability, students must feel that they are performing a "real" communicative act and that they are being evaluated on criteria that do indeed measure the degree of their success in communicating". This test consisted of situations similar to everyday life situations and learners acted in the created situations.
- American English file (book 2), intermediate vocabulary by B J Thomas.
- Camera for recording pretest, posttest.
- Dramatic instruments like plays, stories, films, costumes and masks in order to develop classroom activities.

3.3. The Pattern of Treatment Procedure

As the theory of teaching in drama method is communicative language teaching, the activities used in the present study were: role plays, improvisation, simulation, mime, play reading and play writing. The resources were: text book (American English file 2, intermediate vocabulary), student made materials, movies, videos, music, pictures downloaded from internet. Four chapters of the American English file and 10 chapters of intermediate vocabulary were chosen to be taught. The process in the class was as follows:

- **Review:** Review happened through pantomime or play writing. Students wrote a small play with recent learnt words and structures, at home. Thus, they read that in the class as a narrator by using body language, facial expression, and different voice, or one student pantomimed a word; others guessed and made sentence.
- **Process:** Sometimes teacher prepared a piece of play, a part of movie as introduction. This attracting point was not done just through showing that picture or movie but through asking students to say their sentences about that. So, step by step a story was made. Then students tried to act different characters of this story and make conversations. At last, the dialogues were written. In a different technique, the teacher told a story by showing pictures she had prepared before. Teacher emphasized or repeated the new words or new structures when she was telling the story. After that the list of new words or structures was made. Next step was acting improvisation. They acted out the narrated story. There was the chance for explaining directly and clearly the new grammatical points. New grammar was explained directly and clearly. Then considering that the role play had been recently acted out, the dialogues and sentences were extracted, analyzed and drilled. To help the students internalize the new content, they were asked to think and imagine a new situation, write dialogues with new content and act the story.
- **Evaluation:** the last 10-15 minutes of class was a good time for evaluating students individually. This happened through lecture or play reading. For giving lecture, the student who had been specified for that session gave a lecture, a speech, a recipe or something like that. The students had to choose a different characteristic. They could give the recipe as a professional or an amateur. It was a good activity which helped teacher evaluate the fluency. Another technique was play reading. Students read a play by facial expression, and a little body movement. The play would be from market or student made. Pronunciation was emphasized in this task. Teacher was able to evaluate more precisely the intended individual student.

The treatment consisted of 20 sessions. At the middle of both treatment durations there was midterm exam which consisted of 2 parts: paper and pencil test, practical test. Considering the fact that one lesson had been learnt in previous lessons, teacher showed a picture to the learners and asked them to speak about that picture as an artist. To every student, he showed a different picture. The criteria for evaluation included 4 items: fluency, effort to communicate, quality of communication (grammaticality), comprehensibility.

Some classes were held out of institution to make students try to speak and communicate in a different atmosphere, such as coffee shop, restaurant, museum and bazaar. This brought about new topics and new challenges which are more real-life situation and related to student's needs. Speaking Persian was forbidden from first moment till last moment of these trips. They could use body language, pantomime facial expression to communicate if they faced a problem for speaking English.

3.4. Pre-test and Post-test

All participants in both treatments were pre-tested and post-tested. These tests consisted of situations similar to everyday life situations and learners acted them on. The situations were:

- Shopping in a supermarket. (Appendix B)
- Explaining accident as a witness.
- Reserving and booking flight and a room in hotel in a travel agency.
- Making appointment for visiting doctor.

A teacher played as the contact or participant in conversation (seller, agency manager, and police). The students' performance in the pretest was video recorded, and then it was rated by two raters according to 4 aforementioned criteria. Students knew that they were video recorded.

3.5. Data Analysis

Pearson correlation test was used in the present study for analyzing the ratings given by the two raters to the participants' performances. Normality of data distribution of the control groups and the experimental groups in the pretests and the posttests was assessed through Shapiro-Wilk test. Scores obtained from the first treatment were normally distributed which resulted in using paired t-test and independent t-test. In the second treatment, distribution of the scores was not normal which resulted in using Man-Whitney and Wilcoxon tests. All data was processed through SPSS 16.

4. Results

4.1. First Treatment

Pearson correlation test was used in the present study for analyzing the ratings given by the two raters to the participants' performances. Speaking scores assigned by rater A and rater B were positively correlated, Pearson's $r(10) = .921, .880, .95, .919, p < .01$. High correlation between scores of intended groups assigned by rater A and rater B showed that one set of scores (gained from rater A or B) could be used for testing the hypothesis.

Normality of data distribution of control group and experimental group in the pretest and the posttest was assessed through Shapiro-Wilk test. Scores were distributed normally. Parametric tests were used, including independent t-tests and paired t-test. The mean score of control group in the pre-test was (Mean=11.60, SD=2.79) which increased to (Mean=12.10 SD = 3.03) after treatment. The mean score of experimental group in pre-test was (Mean=11.60, SD= 1.77) which increased to (Mean=15.60, SD= 1.64) after treatment. Moreover, in the case there was any significant difference between mean scores of each group before treatment and after treatment, each group scores in pre- test and posttest should be compared.

Table 1: Results of Paired t-test for Each of Control and Experimental Groups

	Mean	Std. Deviation	SEM	t	df	Sig. (2 tailed)
EXP.PRE - EXP.POST	4.00000	.66667	.21082	-18.974	9	.000
CONT.PRE -CONT.POST	.50000	.97183	.30732	-1.627	9	.138

The paired t-test indicated that no significant difference was observed in control group's speaking scores in pre-test and posttest, because the p value =.13 and $p \geq .05$, but there was a significant difference in experimental group's speaking scores because p value=.00 and $p < .05$. In the next analyses the two groups were compare through the application of an Independent t test.

Table 2: Group Statistics for independent t test

	GROUP	N	Mean	Std. Deviation	Std. Error Mean
PRETEST	1	10	11.6000	1.77639	.56174
	2	10	11.6000	2.79682	.88443
POSTTEST	1	10	15.6000	1.64655	.52068
	2	10	12.1000	3.03498	.95975

Table 3: Results of Independent t-test for Experimental and Control Group in Pre-test and Posttest

		Levene's Test for Equality of Variances		t-test for Equality of Means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
PRETEST of the groups	Equal variances assumed	4.167	.056	.000	18	1.000	.00000
	Equal variances not assumed			.000	15.245	1.000	.00000
POSTTEST of the two groups	Equal variances assumed	4.914	.040	3.205	18	.005	3.50000
	Equal variances not assumed			3.205	13.876	.006	3.50000

The independent t-tests indicated that control group's speaking scores were not significantly different (Mean=11.60, SD=2.79) from experimental group's speaking scores in pre-test (Mean=11.60, SD=1.77), $t(9) = 0.00$, $p < .05$. It means both groups were not statistically different. However independent sample t-test showed that control group's score was highly different (Mean=12.10, SD= 3.03) from experimental group's speaking scores in posttest (Mean=15.60, SD=1.64), $t(9) = 3.205$, $p < 0.05$. Results are shown in Table 3.

4.2. Second Treatment

4.2.1. Fluency

Pearson correlation test was used for analyzing the ratings given by the two raters to the participants' fluency performances. Fluency scores assigned by rater A and rater B were positively correlated, Pearson's $r(15) = .692, .807, .744, .692$, $p < .01$. High correlation between fluency scores of intended groups assigned by rater A and rater B showed that one set of fluency scores (gained from rater A or B) could be used for testing the hypothesis (Appendix C). Normality of data distribution of control group and experimental groups' fluency scores in the pretest and the posttest was assessed through Shapiro-Wilk test. Scores were distributed unimodally. So nonparametric tests were used, including Mann-Whitney and Wilcoxon tests.

The mean score of control group in the pre-test was (Mean=2.60, SD= .632) which after treatment increased to (Mean=2.80, SD= .774). The mean score of experimental group in pre-test was (Mean=2.40, SD= .507) which after treatment increased to (Mean=3.93, SD= .593). Moreover, in case there was any significant difference between mean scores of each group before treatment and after treatment, each group fluency scores in pre- test and posttest would be compared.

Table 4: Result of Wilcoxon Test for Experimental and Control Groups' Fluency Scores in Pre and Posttest

	EX.POST - EX.PRE	CON.POST - CON.PRE
Z	-3.493 ^a	-1.732 ^a
Asymp. Sig. (2-tailed)	.000	.083

The Wilcoxon test result indicated that no significant difference was observed in control group's fluency scores in pre-test and post-test $P = .08$, $p \geq .05$, but there is significant difference in experimental group's fluency scores $P = .00$, $p < .05$. Results are shown in Table 4. Also, in case

there was any significant difference between the mean scores of control group and experimental groups in pre-test and posttest, the mean scores would be compared.

Table 5: Results of Mann-Whitney U for Experimental and Control Group in Pre-test and Post test

	pretest	posttest
Mann-Whitney U	94.500	33.000
Z	-.852	-3.507
Asymp. Sig. (2-tailed)	.394	.000

Exact Sig. [2*(1-tailed Sig.)]

The Mann-Whitney test indicated that control group's fluency scores were not significantly different (Mean=2.60, SD= .632) from experimental group's fluency scores in pre-test (Mean=2.40, SD= .507), $z = -.852$, $p \geq .05$. It means both groups were not statistically different. The Mann-Whitney test showed that experimental group's fluency scores were highly different (Mean=3.93, SD= .593) than control group's scores in posttest (Mean=2.80, SD= .774), $z = -3.507$, $p < 0.05$. Results are shown in Table 5.

4.2.2. Comprehensibility

Pearson correlation test was used for analyzing the ratings given by the two raters to the participants' comprehensibility performances. Comprehensibility scores assigned by rater A and rater B were positively correlated, Pearson's $r(15) = .706, .834, .895, .914$, $p < .01$. High correlation between comprehensibility scores of intended groups assigned by rater A and rater B showed that one set of comprehensibility scores (gained from rater A or B) could be used for testing the hypothesis.

Normality of data distribution of control group and experimental groups' comprehensibility scores in the pretest and the posttest was assessed through Shapir-Wilk test. Scores were distributed unmorally. So nonparametric tests were used, including Mann-Whitney and Wilcoxon tests.

The mean score of control group in the pre-test was (Mean=3.13, SD= .915) which after treatment increased to (Mean=3.53, SD= .915). The mean score of experimental group in pre-test was (Mean=3.53, SD= .639) which after treatment increased to (Mean=4.26, SD= .703). Moreover, in case there was any significant difference between mean scores of each group before treatment and after treatment, each group comprehensibility scores in pre- test and posttest would be compared.

Table 6: Result of Wilcoxon Test for Experimental and Control Groups' Comprehensibility Scores in Pre and Posttest

	EX.POST - EX.PRE	CON.POST - CON.PRE
Z	-3.051 ^a	-1.730 ^a
Asymp. Sig. (2-tailed)	.002	.084

The Wilcoxon test result indicated that no significant difference was observed in control group's comprehensibility scores in pre-test and posttest, $z = -1.730$, $p < .05$, but there is significant difference in experimental group's comprehensibility scores $z = -3.051$, $p < .05$. Results are shown in Table 6. Also, in case there was any significant difference between the mean scores of control group and experimental groups in pre-test and posttest, the mean scores would be compared.

Table 7: Result of Mann-Whitney Test for Experimental and Control Groups' comprehensibility Scores in Pre-test and Posttest

	pretest	posttest
Mann-Whitney U	82.500	63.500
Z	-1.342	-2.223
Asymp. Sig. (2-tailed)	.180	.026
Exact Sig. [2*(1-tailed Sig.)]		

The Mann-Whitney test indicated that control group's comprehensibility scores were not significantly different (Mean=3.13, SD= .915) from experimental group's comprehensibility scores in pre-test (Mean=3.53, SD= .639), $z = -1.342$, $p < .05$. It means both groups were not statistically different. The Mann-Whitney test showed that experimental group's comprehensibility scores were highly different (Mean=4.26, SD= .703) than control group's scores in posttest (Mean=3.53, SD= .915), $z = -2.223$, $p < .05$. Results are shown in Table 7.

4.2.3. Grammaticality

Pearson correlation test was used for analyzing the ratings given by the two raters to the participants' grammaticality performances. Grammaticality scores assigned by rater A and rater B were positively correlated, Pearson's $r(15) = .864, .857, .848, .884$, $p < .01$. High correlation between grammaticality scores of intended groups assigned by rater A and rater B showed that one set of grammaticality scores (gained from rater A or B) could be used for testing the hypothesis. Normality of data distribution of control group and experimental groups' grammaticality scores in the pretest and the posttest was assessed through Shapir-Wilk test. Scores were distributed unimodally. So nonparametric tests were used, including Mann-Whitney and Wilcoxon tests.

The mean score of control group in the pre-test was (Mean=2.86, SD= .833) which after treatment increased to (Mean=2.87, SD= .703). The mean score of experimental group in pre-test was (Mean=3.13, SD= .743) which after treatment increased to (Mean=4.20.26, SD= .774). Moreover, in case there was any significant difference between mean scores of each group before treatment and after treatment, each group grammaticality scores in pre- test and posttest would be compared.

Table 8: Result of Wilcoxon Test for Experimental and Control Groups' Grammaticality Scores in Pre and Posttest

	EX.POST - EX.PRE	CON.POST - CON.PRE
Z	-3.025 ^a	-.577 ^b
Asymp. Sig. (2-tailed)	.002	.564

The Wilcoxon test result indicated that no significant difference was observed in control group's grammaticality scores in pre-test and posttest, $z = -.577$, $p < .05$, but there is significant difference in experimental group's grammaticality scores $z = -3.025$, $p < .05$. Results are shown in Table 8. Also, in case there was any significant difference between the mean scores of control group and experimental groups in pre-test and posttest, the mean scores would be compared.

Table 9: Result of Mann-Whitney Test for Experimental and Control Groups' Scores in Pre and Posttest

	pretest	posttest
Mann-Whitney U	91.500	22.500
Z	-.927	-3.873
Asymp. Sig. (2-tailed)	.354	.000
Exact Sig. [2*(1-tailed Sig.)]		

The Mann-Whitney test indicated that control group's grammaticality scores were not significantly different (Mean=2.86, SD= .833) from experimental group's grammaticality scores in pre-test

(Mean=3.13, SD= .743), $z = -.927$, $p < .05$. It means both groups were not statistically different. The Mann-Whitney test showed that experimental group's grammaticality scores were highly different (Mean=4.20, SD= .774) than control group's scores in posttest (Mean=2.87, SD= .703), $z = -3.873$, $p < .05$. Results are shown in Table 9.

4.2.4. Willingness to Communicate (WTC)

Pearson correlation test was used for analyzing the ratings given by the two raters to the participants' WTC performances. WTC scores assigned by rater A and rater B were positively correlated, Pearson's $r(15) = .700, .732, .767$ at $p < .01$ and $.583$ at $p < .05$. High correlation between WTC scores of intended groups assigned by rater A and rater B showed that one set of WTC scores (gained from rater A or B) could be used for testing the hypothesis. Normality of data distribution of control group and experimental groups' WTC scores in the pretest and the posttest was assessed through Shapir-Wilk test. Scores were distributed unmorally. So nonparametric tests were used, including Mann-Whitney and Wilcoxon tests.

The mean score of control group in the pre-test was (Mean=2.46, SD= .516) which after treatment did not change (Mean=2.46, SD= .516). The mean score of experimental group in pre-test was (Mean=2.46, SD= .516) which after treatment increased to (Mean=3.20, SD= .414). Moreover, in case there was any significant difference between mean scores of each group before treatment and after treatment, each group WTC scores in pre- test and posttest would be compared.

Table 10: Result of Wilcoxon Test for Experimental and Control Groups' WTC Scores in Pre and Posttest

	EX.POST - EX.PRE	CON.POST - CON.PRE
Z	-2.598 ^a	.000 ^b
Asymp. Sig. (2-tailed)	.009	1.000

The Wilcoxon test result indicated that no significant difference was observed in control group's WTC scores in pre-test and posttest, $z = .000$, $p < .05$, but there is significant difference in experimental group's WTC scores $z = -2.598$, $p < .05$. Results are shown in Table 10. Also, in case there was any significant difference between the mean scores of control group and experimental groups in pre-test and posttest, the mean scores would be compared.

Table 11: Result of Mann-Whitney Test for Experimental and Control Groups' WTC scores in Pre and Posttest

	pretest	posttest
Mann-Whitney U	97.500	42.000
Z	-.733	-3.430
Asymp. Sig. (2-tailed)	.464	.001
Exact Sig. [2*(1-tailed Sig.)]		

The Mann-Whitney test indicated that control group's WTC scores were not significantly different (Mean=2.46, SD= .516) from experimental group's WTC scores in pre-test (Mean=2.46, SD= .516), $z = -.733$, $p < .05$. It means both groups were not statistically different. The Mann-Whitney test showed that experimental group's WTC scores were highly different (Mean=3.20, SD= .414) than control group's scores in posttest (Mean=2.46, SD= .516), $z = -3.430$, $p < .05$. Results are shown in Table 11.

5. Discussion and conclusion

In first treatment, results of independent samples t-test revealed a significant difference between the two groups and an improvement in the posttest (Sig = .006). The results of paired t-test revealed that difference between pre-test and post-test in experimental is significant (Sig = .001). In second treatment, results of Mann-Whitney tests revealed significant difference between the experimental and control groups' fluency, comprehensibility, grammaticality and willingness to communicate in posttests (sig =.000, .026, 000, and .001_respectively). The results of Wilcoxon tests revealed that

difference between pre-tests and posttests in experimental group's fluency, comprehensibility, grammaticality and effort to communicate is significant (sig = .000, .002, .002, and .009 respectively).

In order to figure out the reasons for such findings it is better to review relative claims. Drama method is useful because:

- It is an ideal way to bring skills of grammar, reading, writing, speaking, listening together where the focus is not on the form but on fluency and meaning (Dodson, 2000, p.139).
- Through these interactions, learners figure out what they need to do to keep the conversation going and make themselves comprehensible (Lightbown and Spada, 2002, p.43). Via (1987) defined drama techniques as strategies to communicate or convey the intended meaning and make one's comprehensible which includes a wide range of activities (Cited in Janudom & Wasanasomsithi, 2009, p. 4).
- It is stated that variables like [...] , and motivation influence on WTC (Donovan & MacIntyre, 2004; Hashimoto, 2007; MacIntyre, 1994) and on the other hand Dougill (1987) and Teylor (2000) sated that drama techniques can satisfy the needs of language learners in that they create motivation, enhance confidence, and provide context in learning a language (cited in Janudom & Wasanasomsithi, 2009, p. 4).
- Drama method reinforces the need to speak so increases the willingness to communicate (Wessels, 1987, cited in in Janudom & Wasanasomsithi, 2009, p. 4)
- Nunan (1999) claimed that while grammar lessons can work for some people but opportunity to use the language can benefit all the learners. Dewey (1938) said formal instruction should incorporate activities of language use so that their learners can put theories into practice (Wu, C. H., 2007, p.62)
- Teachers by using drama method as their instruction method can provide learners with more interaction, more self-confidence, less anxiety, more motivation, and more imagination as well as opportunity to practice meaningfully what they learn so they most probably improve fluency, grammaticality, effort to communicate as well as comprehensibility of language learners. Syllabus designers and material developers can examine EFL learners' and teachers' needs considering the use of drama-based instruction method and activities. This is because lack of materials which facilitate using this method such as, textbooks, play books, movies, and media player equipment is really felt in Iran.

References

- Adebisi, A., & Adelabu. B. (2014). Improvisations as a tool for developing students' competence in English language: a study of the federal university of agriculture, Abeokuta, Nigeria. *International Journal of Education and Research*, 1(11).
- Bartz, W. H. (1979). Testing oral communication in the foreign language classroom. *Arlington: Center for applied linguistics*.
- Brown, J. D. (1996). Fluency development. In G. van Troyer (Ed.) *JALT '95: Curriculum and evaluation*, 174-179. Tokyo: Japan Association for Language Teaching
- Chauhan, V. (2004). Drama techniques for teaching English. *The Internet TESL Journal*, 10(10).
- Bolton, G. M. (1984). *Drama as education: An argument for placing drama at the centre of the curriculum*. Addison-Wesley Longman Ltd.
- Dodson, S. L. (2000). FAQs: Learning languages through drama. *Texas Papers in Foreign Language Education*, 5(1), 129-141.
- Donovan, L.A., and MacIntyre, P.D. (2004). Age and sex differences in willingness to communicate, communication apprehension, and self-perceived competence, *Communication Research Reports*, 21,420- 427.

- Dora To, L., Phoebe Chan, Y., Lam, Y. K., & Tsang, S. Y. (2011). Reflections on a primary school teacher professional development program on learning English through Process Drama. *Research in Drama Education: The Journal of Applied Theatre and Performance*, 16(4), 517–539. Retrieved from <https://doi.org>
- Fillmore, C. J., Kempler, D., & Wang, W. S. (Eds.). (2014). *Individual differences in language ability and language behavior*. New York: Academic Press.
- Gatbonton, E., & Segalowitz, N. (1988). Creative automatization: principles for promoting fluency within a communicative framework. *TESOL Quarterly*, 22, 473-492.
- Gorjian, B., Moosavinia, S., & Jabripour, A. (2010). Dramatic performance in teaching drama in EFL context. *TESL-EJ*, 13(4), 1-13.
- Hashimoto, Y. (2002). Motivation and willingness to communicate as predictors of reported L2 use: the Japanese ESL context. *Second Language Studies*, 20, 29-70.
- Janudom, R., & Wasanasomsithi, P. (2009). Drama and questioning techniques: Powerful tools for the enhancement of students' speaking abilities and positive attitudes towards EFL learning. *ESP World*, 8(5), 23-28.
- Kellem, H. (2009). The form-meaning response approach: poetry in the EFL classroom. *English Teaching Forum*, 47(4), pp. 12-17.
- Ladousse, G.P. (1987). *Role play*. New York: Oxford University Press.
- Mahmoodi, M. H., & Moazam, I. (2014). Willingness to communicate (WTC) and L2 achievement: The case of Arabic language learners. *Procedia-Social and Behavioral Sciences*, 98, 1069-1076.
- MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, 11, 135-142.
- MacIntyre, P. D., Dornyei, Z., Clement, R., & Noels, k. A. (1998). Conceptualizing willingness to communicate in an L2: Situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545- 562.
- Miccoli, L. (2003). English through drama for oral skills development. *ELT journal*, 57(2), 122-129.
- Nunan, D. (2001). *Second Language Teaching and Learning*. Boston, MA: Heinle and Heinle.
- Shen, L., & Suwanthep, J. (2011). E-learning Constructive Role Plays for EFL Learners in China's Tertiary Education. *Online Submission*, 49.
- Suganda, L.; Zuraida, Z. and Kurniawan, D. (2017). The Effect of Role-Playing for Building Students' Speaking Motivation and Positive Characters. *In the Tenth Conference on Applied Linguistics and The Second English Language Teaching and Technology Conference in collaboration with The First International Conference on Language, Literature, Culture, and Education*, 1, 295-298.
- Lightbown, P. M., & Spada, N. (2002). *How languages are learned*. Oxford: Oxford University Press.
- Vasanthrao, P. P. (2012). Communicative approach: Its application and relevance in 21st century. *The Criterion: An International Journal in English*, 3(4), 2-9.
- Sam, Wan (1990). Drama in teaching English as a second language: A communicative approach. *The English Teacher*, 9, 1-11
- Wessels, C. (1987). *Drama*. Oxford: Oxford University Press.
- Wu, C. H. (2007). *Spoken grammaticality and EFL teacher candidates: measuring the effects of an explicit grammar teaching method on the oral grammatical performance of teacher candidates* (PhD Dissertation). Retrieved from <https://etd.ohiolink.edu>

Appendix A: Criteria for Rating Pre-test and Post-test (Bartz, 1979)

FLUENCY

General definition: Fluency does not refer to absolute speed of delivery, since native speakers of any language often show wide variations in this area. Fluency refers to overall smoothness, continuity, and naturalness of the student's speech, as opposed to pauses for rephrasing sentences, grouping for words, and so forth.

Definition of each level on the scale:

1. Very many unnatural pauses, very halting/ and fragmentary delivery
2. Quite a few unnatural pauses, frequently halting and fragmentary delivery
3. Some unnatural pauses occasionally halting and fragmentary delivery
4. Hardly any unnatural pauses, fairly smooth and effortless delivery
5. No unnatural pauses, almost effortless and smooth, but still perceptibly non-native
6. As effortless and smooth as speech of native speaker

1	2	3	4	5	6
----------	----------	----------	----------	----------	----------

COMPREHENSIBILITY

General definition: The ability of the student to make himself understood, to convey meaning

Definition of each level on the scale:

1. No comprehension, couldn't understand a thing student said
2. Comprehended small bits and pieces, isolated 'words
3. Comprehended some phrases or word clusters
4. Comprehended short, simple sentences
5. Comprehended most of what student said
6. Comprehended all of what student said

1	2	3	4	5	6
----------	----------	----------	----------	----------	----------

QUALITY OF COMMUNICATION (Grammaticality)

General definition: The grammatical correctness of the student's utterances

Definition of each level on the scale:

1. No utterances rendered correctly
2. Structure of very few utterances rendered correctly
3. Some utterances rendered correctly, but many structural problems remain
4. Many correct utterances, but some problems remain with structures
5. Most utterances rendered correctly; only minor problems with structure
6. All utterances rendered correctly

1	2	3	4	5	6
----------	----------	----------	----------	----------	----------

EFFORT TO COMMUNICATE (willingness to communicate)

General definition: The student's willingness to express himself and to get his message across. How hard does the student try to make himself understood? Does he make any attempt to express himself? Does he use gestures to help express himself? Or does he withdraw into an embarrassed silence that makes it very difficult for him to communicate at all? Ask yourself this question: To what degree does the student show an effort and a willingness to express himself in target language?

Definition of each level on the scale:

1. Student makes little effort to communicate, doesn't seem to care if he completes the task
2. Student makes some effort to communicate but does not try very hard to complete the task
3. Student makes an effort to communicate, tries to complete the task, and may add something not required by the task

4. Student makes a real effort to communicate, tries very hard to complete the task, and may add something not required by the task
5. Student makes a special effort to communicate, shows an extremely high effort to complete the task, and goes beyond the required task
6. Student makes an unusually high effort to communicate, shows an almost overzealous effort to complete the task, goes way, beyond the required task);, and uses all possible resources, verbal and non-verbal, to express himself

1	2	3	4	5	6
---	---	---	---	---	---

Appendix B: A sample of situation for role-play in tests

SHOPPING IN A FOOD MARKET

Preparation: The teacher prepares a set of index cards on which pictures of grocery items are drawn or pasted.

Instructions to the Student: You are in an English language country food market. You have completed all your shopping except one item, for which you cannot remember the (foreign language) word. A picture of the item is on the index card you have selected from a pack of cards. Describe the item to the "clerk" (teacher) so that he or she will be able to find it. (The teacher cannot see the card that the student is describing and must try to select the correct item based on the student's description)

Appendix C: A sample of learners; scores

Students' fluency score in pre/post test; Second treatment

Experimental group Pretest	Experimenta group Posttest	Control group Pretest	Control group Posttest
2.0	2.0	4.0	2.0
2.0	4.0	4.0	4.0
3.0	2.0	4.0	2.0
2.0	3.0	3.0	4.0
3.0	2.0	5.0	2.0
3.0	2.0	4.0	2.0
3.0	3.0	4.0	3.0
2.0	3.0	3.0	3.0
2.0	3.0	4.0	3.0
2.0	3.0	4.0	3.0
3.0	3.0	4.0	4.0
3.0	2.0	4.0	2.0
2.0	2.0	4.0	2.0
2.0	2.0	3.0	3.0
2.0	3.0	5.0	3.0

Students' comprehensibility scores in pre/post test: second treatment

Experimental group pretest	Control group pretest	Experimental group posttest	Control group posttest
4.0	2.0	4.0	2.0
4.0	4.0	6.0	4.0
4.0	3.0	5.0	3.0
3.0	5.0	4.0	5.0
5.0	3.0	5.0	3.0
4.0	3.0	5.0	3.0
4.0	3.0	4.0	4.0
3.0	4.0	4.0	4.0
3.0	4.0	4.0	4.0
3.0	4.0	4.0	4.0
3.0	3.0	4.0	2.0
3.0	2.0	4.0	4.0
4.0	2.0	4.0	3.0
3.0	3.0	4.0	5.0
3.0	2.0	3.0	3.0