

INVESTIGATING THE ROLE OF FORMATIVE MOBILE BASED ASSESSMENT IN VOCABULARY LEARNING OF PRE-INTERMEDIATE EFL LEARNERS IN COMPARISON WITH PAPER BASED ASSESSMENT

Dr. Nahid YARAHMADZEHI

ORCID: 0000-0003-4601-0557
English Language Department
Management and Humanities Faculty
Chabahar Maritime University
Chabahar, IRAN

Mostafa GOODARZI

ORCID: 0000-0002-4076-3147
English Language Department
Management and Humanities Faculty
Chabahar Maritime University
Chabahar, IRAN

Received: 20/02/2019 **Accepted:** 21/05/2019

ABSTRACT

Throughout this study technology and especially mobile phones was utilized in EFL classrooms in order to see whether it can influence the process of vocabulary formative assessment and consequently improve vocabulary learning of Iranian pre-intermediate EFL learners or not. Two groups of pre-intermediate EFL learners participated in this study. Regarding the first group (FMA) the vocabulary learning of learners was assessed formatively during ten sessions using Socrative mobile application. The vocabulary learning of the second group (FPA) was also assessed formatively but paper and pen were the instruments used by instructor and students to take the tests. After applying quasi-experimental research design including pretest, treatment and posttest and after running independent sample t tests to posttest scores, the results showed that those pre-intermediate EFL learners attending in the group where vocabulary gain was assessed formatively using mobile application named Socrative performed significantly better in posttest than group assessed formatively based on paper and pen ($Sig=0.03<0.05$). The analysis of attitude questionnaire distributed among participants of the group assessed formatively using mobile phone exhibited that they possessed a positive attitude towards mobile based testing.

Keywords: Formative assessment, mobile based assessment, paper based assessment, attitude, vocabulary, pre-intermediate EFL learners

INTRODUCTION

Background

Since computers and smartphones have become essential elements of human life in many areas like business, offices, and homes, language teachers have also tried to apply such technology as a new pedagogical tool in their classrooms in order to improve their students' language learning. As a result, this positive outlook towards computers and smart phones' usage in language learning has caused the emergence of computer assisted language learning (CALL) and mobile assisted language learning (MALL) methods. CALL and MALL are widely used to refer to the area of technology and its contributions to second language teaching and learning (Chapelle, 2001).

Nowadays, because of their flexibility, size and capacities, mobile phones are beginning to replace computers. Consequently, many researchers believe that mobile phones are not only able to support formal and informal

learning but also to complete the process of learning via computers. As Kukulska-Hulme, Norris and Donohue (2015) stated, students nowadays carry with them powerful devices with which they can:

- create and share multimodal texts
- communicate spontaneously with people anywhere in the world
- capture language use outside the classroom
- analyze their own language production and learning needs
- construct artifacts and share them with others
- provide evidence of progress gathered across a range of settings, in a variety of media.

Since Chickering and Ehrmann (1996) coined the term MALL (Mobile Assisted Language Learning), mobile phones have recorded a tremendous growth in language learning and teaching field. In fact, MALL refers to the language learning that is assisted or enhanced through the use of a handheld mobile device (Chinnery 2006) and it is a subdivision of Computer Assisted Language Learning (Yang 2013). Therefore, many research projects started to investigate the role of MALL in various areas of language learning and teaching such as basic language skills including listening (Azara and Nasiri, 2014), reading comprehension (Chen & Hsu, 2008; Plana, Gimeno & Appel, 2013), pronunciation and speaking performance (Ducate & Lomicka, 2009; Han & Keskin, 2016; Mahmoud, 2013), grammatical accuracy (Baleghizadeh & Oladrostam, 2010) and writing (Allagui, 2014; Mahmoud, 2013) as well as subskills namely vocabulary (Taki and Khazaei, 2011) and idioms (Sahan, Coban & Razi, 2016).

One area of language learning and teaching classes which can be influenced by technology usage is assessment. Bachman (2004) defined assessment as “a process of collecting information about something that we are interested in, according to procedures that are systematic and substantially grounded” (pp. 6-7). The result of an assessment procedure can be a score or a verbal description. Actually, teaching and learning processes always consist of two main essential and inseparable activities, namely: teaching and assessment. The two cannot be separated since teaching is always followed by administering an assessment to get an insight into the students’ ability and their understanding of the learning material which is being taught.

This study aimed to work on the role of mobile based testing in an EFL environment. Using computers and online processes are vastly investigated before and the advantages and disadvantages of the role of computers as one area of technology in testing have been discussed. This study, thus, aimed to investigate any difference in learners’ performance in mobile based tests compared with paper based ones. As a matter of fact, a growing number of students possess smartphones or tablets and it is beneficial to guide learners towards an effective usage of such technology instead of forbidding them from bringing their smartphones to class.

On the other hand, the application of mobile phones in formative assessment is investigated in this study. Formative assessment or assessment for learning takes place during learning rather than after learning which refers to summative assessment, and has an entirely different function: to provide ongoing feedback to both students, to improve their learning, and teachers, to improve their teaching (Spector et al., 2016). Thus beyond a difference in purpose, there are also crucial differences in timing between the summative and formative assessments: formative testing results are especially useful when they become available early in the learning process. Thus, this study aimed to investigate the role of smartphones in vocabulary learning of English students in a formative way.

Research Questions

The research question and related hypothesis is in the following:

- Q1: Is there any significant difference between the vocabulary gain of Iranian pre-intermediate EFL learners assessed formatively by paper and pen and those assessed formatively using a mobile device?
- Q2: What is the attitude of Iranian pre-intermediate EFL learners towards mobile based assessment?

REVIEW OF LITERATURE

Assessment

Assessment could be defined as a systematic process of gathering information about what a student knows, is able to do, and is learning to do. The information obtained from assessment, on the other hand provides a foundation for decision-making and planning for further instruction and learning. Therefore, assessment is an integral part of instruction that enhances, empowers, and celebrates student learning. However, testing is a way of conducting assessment which is technically associated with definite timing and settled procedures (Brown, 2004). Briefly, by using a variety of assessment techniques, teachers gather information about what students know and are able to do, and provide positive and supportive feedback to students.

Summative vs. Formative Assessment

Another useful distinction to bear in mind is the function of an assessment: How is the procedure to be used? Two functions are commonly identified in the literature: formative versus summative assessments. Summative assessments are used to measure learning. They are typically administered at the end of an academic semester or year to discern whether and to what extent students have mastered concepts taught during the course; thus, summative assessments are primarily evaluative in their purpose (Varier, 2014). This type of assessment is also considered a high-stakes assessment because key decisions such as course grades, admissions, promotions, and evaluations of teacher and school performance are based on how students fare on these tests. Examples of summative assessments include standardized annual examinations in schools and mid-term and final examinations in college courses.

Formative assessment, on the other hand, refers to a process in which assessments are used to inform instruction and support student learning. They are used by teachers to make instructional decisions, and by students, to become aware of their progress. Formative assessments can be formal in that they serve the diagnostic purpose of identifying gaps in understanding or informal as they are embedded in classroom activities to elicit student understanding. In other words, formative assessment is defined by Brown (2003) as evaluating students in the process of forming their competencies and skills with the goal of helping them to continue that growth process. The key to such formation is the delivery (by the teacher) and internalization (by the student) of appropriate feedback on performance, with an eye toward the future continuation (or formation) of learning.

Vocabulary Teaching and Assessment

Actually, vocabulary is the flesh of a language whereas grammar is the skeleton. In order to be able to use the language productively, students must know certain amount of vocabulary, not only for communicating orally, but also in writing. Nation (2005) regarding the importance of vocabulary believed that vocabulary knowledge enables language use, language use enables the increase of vocabulary knowledge; knowledge of the world enables the increase of vocabulary knowledge and language use and so on. This contextualized approach to learning vocabulary can definitely help students increase their vocabulary through authentic interaction.

According to Laufer (1997), among the components of language, vocabulary learning is a language component that has been the study matter of many linguists for many years. Probably because many scholars believe that vocabulary learning is at the heart of language learning and language use. In fact, it makes the essence of a language. Hence, as Laufer (1986) claimed, speakers cannot convey meaning and communicate with each other in a particular language without sufficient vocabulary knowledge.

Assessing the vocabulary knowledge of EFL learners is both essential and reasonably straightforward. It is essential in the sense that words are considered to be the fundamental blocks of language, the units of meaning from which larger structures, such as phrases, sentences, paragraphs and texts are drawn. Many language teachers and applied researchers now recognize the importance of vocabulary learning and investigate ways of promoting it more efficiently. Hence, from different points of view, vocabulary can be seen as a central area in language teaching, requiring testing tools to check learners' progress in vocabulary learning and to assess how sufficient their vocabulary knowledge is in order to meet their communicative needs (Read, 2000).

Formative Assessment of Vocabulary

Formative assessment is anything the teacher does to assess or test her students' levels of understanding about a subject while they are still learning that subject. There is a large range of formative assessment methods and techniques available to be implemented in the vocabulary learning process. It may be a quiz or a test, an oral question-and-answer session or a one-sentence summary, a one-minute paper or a group presentation. Using formative assessment and feedback gives students time to pause and evaluate their performances before they reach major assessments.

Therefore, formative assessment plays a key role in vocabulary learning within the field of foreign language learning. Read (2000) claimed that 'discrete, selective, context-independent vocabulary tests have been an integral part of the educational measurements science for the whole twentieth century' (p. 115). They possess all the virtues of an objective language test and for a long time were considered to be very productive. Tests such as multiple-choice, gap-filling, matching, true/false items, cloze tests and C-tests are still greatly used.

Mobile Based Assessment

Mobile-Based Assessment (MBA) is a relatively new mode of assessment that is delivered through mobile devices (PDAs, mobile phones, smartphones, tablets). MBA can be used both in the context of a pure mobile learning approach and also in a blended learning approach, as complementary or alternative to computer- or paper-based assessment. Mobile devices can be used in different assessment types, either inside or outside the classroom boundaries for formative assessment, self- and peer-assessment, work-based assessment, classroom polling, etc. Mobile devices and wireless technologies offer many affordances that can not only enhance existing but also introduce new opportunities to assess learning. Some of these affordances are: adaptivity and personalization, context-awareness, seamlessness and ubiquity, social media connectivity. One main advantage of using mobile applications for testing is the immediate feedback provided for both teacher and students. Actually, in order to have a fruitful formative assessment, feedback must be given to students in appropriate time. Hence, teachers need time to give feedback in formative assessment based on paper and pen, but using mobile applications for formative assessment, provides both teachers and learners with immediate feedback which can help learners be aware of their strengths and weaknesses in order to improve themselves for future tests and consequently learn better.

All in all, to overcome possible barriers with applying formative assessment, using mobile applications which are designed for testing is going to be investigated in this study. Consequently, in FMA group studied in this project, formative mobile based assessment is allocated and it is compared with another group where formative assessment is based on paper and pen.

METHOD

Participants

The participants of this study were 40 EFL learners who studied General English language at Chabahar Maritime University, Iran. Their age ranged from 18 to 22. The participants had been assigned to two different groups, each with 20 students, prior to the study and hence they were not randomly selected. The participants were all pre-intermediate EFL learners. They were all native speakers of Persian language and Balouchi. The participants of these two groups attended in their General English class 3 hours a week with the same instructor and textbook. The only difference between these groups was the type of assessment allocated to each one. A group assessed formatively using mobile application was named FMA group (n=20). The other class as FPA group (n=20), was also assessed formatively but based on conventional paper and pen tests.

Design

Based on the purpose and the nature of this study, quasi experimental design consisting of a pretest, treatment and a posttest was used to answer the research question. After applying a Longman proficiency test to all participants signifying that they all were at pre-intermediate level of English proficiency, one vocabulary

pretest exhibited that there was no significant difference of vocabulary knowledge among the participants at the outset of the study. The treatment in two groups took 10 sessions to be performed. Since the study was around formative assessment, 10 quizzes based on vocabulary introduced in their textbook as new words were applied. One posttest, then, was designed to examine what two groups learned in these 10 sessions and determined whether there was any significant difference between two groups after treatment or not.

Procedure

After selecting the participants of the study and being assured that all were pre-intermediate EFL learners, the procedure was followed as below:

- **Pretest:** A pretest, an English Vocabulary in Use Pre-intermediate and Intermediate Level Test, was applied to both groups to ensure that they had no significant differences at the outset of the study. Since the only independent variable was the mode of assessment (mobile based or paper based) and vocabulary knowledge was regarded as only dependent variable, an independent samples t-test was run by SPSS software to pretest scores of two groups.
- **Treatment:** After the pretest, the instructor started to teach both classes using the same method. The only difference was the type of assessment used by the teacher. In the group named FMA, formative mobile based assessment was applied. Therefore, during the semester students of this group took 10 vocabulary quizzes, one at the end of each session. These quizzes were based on the words taught in each given session and the sessions before. All the tests included 20 multiple choice items. The other group named FPA was also assessed through formative assessment. The same quizzes with the same questions applied to first group was applied in this group too, but paper and pen were the only instruments used by students to answer the questions.
- **Posttest:** The posttest composed of 20 multiple choice items based on whatever learners learned in their textbook throughout the course, was designed and administered in two groups to compare the scores of two groups after the treatment. Another independent sample t-test was run to posttest scores because the only independent variable was the mode of assessment (mobile based or paper based) and the only dependent variable was vocabulary learning.
- **Attitude:** As the last step in this study, an attitude questionnaire was distributed among participants in FMA group to find out the attitude of learners towards mobile based testing. This questionnaire “the attitudes towards mobile based test” was developed from Computer Anxiety Subscale (Loyd & Gressard, 1984).

Instruments

Longman Proficiency Test

The Longman proficiency test used in this study was provided by Pearson Education Palsky for Catholic university of Lublin in 2004 which consisted of one hundred questions related to grammar and vocabulary. An assortment of guidelines was provided at the end of the questions specifying the range of scores falling within each level of proficiency, indicating that for example those who scored from 20 to 40 belonged to pre-intermediate language learners. The students had 60 minutes to answer all the questions. After correcting the papers and receiving the results, those whose scores were not in the range of pre-intermediate learners were removed from the study. Determining the proficiency level of students helped in composition of more homogeneous groups of participants and in providing an appropriate vocabulary pretest.

Pretest

The pretest of the current study which was applied to all three groups to ensure that they had no significant differences in vocabulary knowledge at the beginning of the study was based on fifty vocabulary multiple choice items of English Vocabulary in Use Pre-intermediate and Intermediate Level Test published by Cambridge Press and designed by Stuart Redman and Ruth Gairns (2008). The results of this pretest assured the researcher that all participants were at the same level of vocabulary knowledge at the outset of the treatment.

Treatment Tests

Ten tests were administered in FMA and FPA groups as formative assessment. All the tests were based on the vocabulary introduced in the textbook taught by the instructor named Concepts and Comments. Each test included 10 to 20 vocabulary multiple choice items. The tests were based on the lessons taught during the present session and all other previous sessions. All the items in two groups were the same and the mode of assessment (mobile versus paper) was the only difference between two groups. Besides, all the tests were organized by the researcher and approved by the instructor. For all treatment tests a Kr_21 formula was used to approve the reliability of the tests.

Posttest

The posttest as one main stage of the study which examined any significant difference in vocabulary learning between all three groups after treatment was based on the new words introduced in each lesson of textbook named Concepts and Comments taught in these groups by the same instructor. It included 20 multiple choice items. The test was designed by researcher and approved by the instructor. To establish a valid posttest only new vocabulary introduced in the textbook was tested. The instructor taught 12 lessons of the textbook and the posttest was also organized based on these twelve lessons leading to the content validity of the posttest. KR-21 formula was used to confirm that the posttest was reliable.

Attitude Questionnaire

As the last instrument used in this study, an attitude questionnaire was distributed among participants in FMA group. This questionnaire “the attitudes towards mobile based test” was developed from Computer Anxiety Subscale, (Loyd & Gressard, 1984). First of all, since the research focused on using mobiles for testing, all statements were modified using the word “mobile” instead of “computer”. Then, 10 items were modified to reflect mobile attitude of those students who took part in vocabulary mobile based tests. The Hoyt estimate of internal consistency reliability of the revised subscale was 0.82 for this sample. The items were also divided into two positive (1_5) and negative (6_10) parts. To be easier for our participants, it was translated into Persian.

Socrative Application

As it was mentioned before the only difference regarding the treatment in this study was the type of assessment applied for each group of language learners. Assessment in FMA group was carried out by a mobile application designed for testing named Socrative by Master Connect. The teacher then creates a virtual room with a particular code in the application designed for teachers (which can be received from website with the address www.socrative.com) and writes his questions there. Then, students install student` App and use the same code to enter the virtual room of questions created and organized by the teacher and answer the questions. The teacher can also manage the way the learners are to answer the questions; whether they should answer all questions at once or one by one. After the test, the teacher receives the results and can immediately share it with the learners. Therefore, throughout this study and to deliver the test onto the smart devices that students employ in class, Socrative Application was chosen as the quiz delivery system. It is a widely used, free web service that allows students to do quizzes or even answer quick questions in a manner of polling by using their smart phones. The students were trained on how to navigate the website/app a week prior to the beginning of Mobile based assessment, however, the training was minimal since no registration was required to use Socrative; students just needed an instructor generated room number.

Since Socrative is the primary tool of this study that represents the Mobile Based-Testing, it is necessary to have a brief understanding of its functions. Diechman (2014) provided the following description:

Socrative 2.0 is an online assessment and student response tool that can be run on any platform that has a connection to the Internet. An AASL Best Website and Best App for 2013, this system was created by a team that is passionate about education, and that passion is obvious in their work. The system can be used as a student engagement tool and as either a formative or summative assessment

mechanism. Laptop or computer users can just log in to join their class through a student website. An app is also available for tablets and smartphones. Two separate apps are needed to use the Socrative 2.0 tool: a Teacher App and a Student App that must be downloaded onto each device. As of this writing, both the online website and the tablet apps are free to all users. (p. 72)

Socrative application was chosen for this study because of its ease of use, its lack of cost (it is a free service), and its ability to quickly assess prior knowledge where students can easily get involved and assess their understanding of lessons; lessons which in this study were included in Concepts and Comments textbook.

DATA ANALYSIS AND RESULTS

Longman Proficiency Test

First of all a Longman proficiency test was administered to ensure that the participants of both groups were pre-intermediate English learners. The assortment proposed with the questions showed the range of scores falling in each level, therefore those who scored from 26 to 45 belonged to pre-intermediate language learners. The results of the proficiency test applied to both groups are shown in table 4.1:

Table 1. Longman proficiency test for two groups

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--------------------|----|---------|---------|---------|----------------|
| FMA | 20 | 19.00 | 62.00 | 33.2500 | 13.43552 |
| FPA | 20 | 12.00 | 54.00 | 30.9000 | 11.81836 |
| Valid N (listwise) | 20 | | | | |

As it is shown in table 4.1, the mean for each group is between 26 and 45 which means that all participants in both FMA and FPA groups could be regarded as pre-intermediate EFL learners. The results showed that the study was going to be conducted dealing with homogeneous learners and their proficiency level had no undesirable effect on the process of data collection and analysis.

Data Analysis for First Question of the Study

Pretest

To answer the first question of the study, a pretest, first of all, was administered to both FMA and FPA groups. The experimental group was labeled as FMA because the vocabulary gain of this class was decided to be assessed formatively using mobile phone. The other group, on the other hand in which vocabulary gain was going to be assessed formatively based on paper and pen was labeled as FPA. The pretest was administered at the outset of the study to confirm that learners were at the same level of vocabulary knowledge before starting the treatment of the study.

Descriptive statistics based on the results of pretest are shown in table 4.2.

Table 2. Descriptive statistics for pretest scores

| | Groups | N | Mean | Std. Deviation | Std. Error Mean |
|------------|--------|----|---------|----------------|-----------------|
| vocabulary | FMA | 20 | 18.6000 | 10.34357 | 2.31289 |
| | FPA | 20 | 19.9000 | 9.55262 | 2.13603 |

A KR-21 formula was applied for analyzing the reliability of the pretest:

$$[n/(n-1) * [1-(M*(n-M)/(n*Var))] = [(50/49)*[1-(18.6*(50-18.6)/50*106.09)]=0.86$$

The reliability of 0.86 declares that the pretest was benefited from a high range of reliability.

Since the only independent variable regarding the first question of the study was the mode of assessment (mobile based or paper based) and the only dependent variable was vocabulary knowledge of participants, an independent sample t-test was run the results are shown in table 4.3.

Table 3. Independent samples T-Test for pretest between FMA and FPA

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|--------------------|-----------------------------|---|------|------------------------------|--------|-----------------|-----------------|-----------------------|---|---------|
| | | F | Sig. | t | Df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | Lower | Upper | |
| Vocabulary Pretest | Equal variances assumed | .255 | | .617 | 38 | .682 | -1.30000 | 3.14835 | -7.67350 | 5.07350 |
| | Equal variances not assumed | | | | 37.762 | .682 | -1.30000 | 3.14835 | -7.67482 | 5.07482 |

The tables 4.3 gives the fact that although mean score of FPA group was higher (19.90>18.60), there was no statistically significant difference between two groups before starting the treatment (sig= 0.617>0.05). Therefore, the vocabulary knowledge of participant was not as much different to affect the process of research.

Reliability of Treatment Tests

Since the study is around formative assessment of vocabulary, 10 quizzes were administered during the term among both groups. All the tests consisting 20 multiple choice items were based on the vocabulary taught during the term. In order to get assured of the reliability of the tests, the scores obtained by participants of FMA group were analyzed. KR_21 formula was employed for measuring the reliability of each test. Table 4.4 shows mean, standard deviation and reliability of each test.

Table 4. Reliability of treatment tests

| Tests | Mean | Standard Deviation | Reliability |
|---------|-------|--------------------|-------------|
| Test 1 | 14.60 | 4.50 | 0.85 |
| Test 2 | 13.05 | 5.04 | 0.87 |
| Test 3 | 14.20 | 4.6 | 0.84 |
| Test 4 | 15.6 | 3.6 | 0.8 |
| Test 5 | 16.3 | 3.7 | 0.82 |
| Test 6 | 14.25 | 3.7 | 0.73 |
| Test 7 | 15.65 | 3.71 | 0.79 |
| Test 8 | 17.45 | 2.16 | 0.70 |
| Test9 | 17.55 | 3.51 | 0.86 |
| Test 10 | 18.20 | 3.27 | 0.89 |

As it is shown in table 4.4, all the treatment tests which were administered through formative assessment were reliable tests.

Posttest

After applying different types of treatment in both groups, to find out any significant difference between vocabulary gains of two groups after the treatment, a posttest was administered. Descriptive statistics regarding the posttest scores are shown in table 4.5.

Table 5. Descriptive statistics for posttest of FMA and FPA

| | Groups | N | Mean | Std. Deviation | Std. Error Mean |
|----------------|--------|----|---------|----------------|-----------------|
| vocab_posttest | FMA | 20 | 83.2500 | 14.62469 | 3.27018 |
| | FPA | 20 | 70.7500 | 21.47673 | 4.80234 |

The reliability of posttest was analyzed employing KR-21 formula:

$$[n/(n-1) * [1-(M*(n-M)/(n*Var))]] = [1.01*1-83.25(100-83.25)/100*21374.44] = 0.94$$

The posttest can also be regarded as a reliable test because the obtained reliability was 0.94.

Since the mode of vocabulary assessment (mobile based or paper based) was the only independent variable beside vocabulary gain of learners as the dependent variable, an independent sample t-test was proposed to be employed to see whether there was a significant difference in vocabulary gain of FMA group and FPA group after applying different modes of assessment throughout the treatment or not; table 4.6 is provided by SPSS software to answer the first question of the study.

Table 6. Independent samples test for posttest between FMA and FPA

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|----------------|-----------------------------|---|------|------------------------------|--------|-----------------|-----------------|-----------------------|---|----------|
| | | F | Sig. | t | Df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | Lower | Upper | |
| vocab_posttest | Equal variances assumed | 4.576 | .039 | 2.151 | 38 | .038 | 12.50000 | 5.81004 | .73819 | 24.26181 |
| | Equal variances not assumed | | | 2.151 | 33.502 | .039 | 12.50000 | 5.81004 | .68611 | 24.31389 |

As it is shown in 4.4 and 4.5, beside the fact that mean score of FMA group where vocabulary formative assessment was based on mobile application was higher than FPA group where vocabulary formative assessment was based on paper (FMA mean=83.25 > FPA mean=70.75), there was also statistically significant difference between two groups after the treatment (sig=0.03 > 0.05). Therefore, the null hypothesis that says there is no significant difference between the vocabulary gain of Iranian pre-intermediate EFL learners tested by paper and pen and those who are tested using mobile devices was **rejected**.

Data Analysis for Second Question of the Study

Quantitative Analysis

The second research question of this study which was about the attitude of pre-intermediate English learners towards mobile based assessment was answered quantitatively and qualitatively. An attitude questionnaire containing ten items was distributed among participants in FMA group. All 20 students of FMA group and 5 other students whose scores were not included in the analyses for first question of the study answered the questions. Descriptive statistics regarding the questionnaire result was run to determine whether the attitude of students towards such type of assessment was positive or negative.

To answer the questions, respondents were asked to choose from a six-point Likert-type scale of agreement running from (1= Strongly Disagree, 2= Disagree, 3= somewhat disagree, 4= somewhat Agree, 5= Agree and six= strongly agree). The results are shown in Table 4.6.

Table 7. Descriptive statistics of questionnaire

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---|----|---------|---------|--------|----------------|
| 1 Using the mobile for taking a test did not scare me at all. | 25 | 2.00 | 6.00 | 5.3200 | 1.24900 |
| 2 I would not be threatened even if my classmates liked taking the test using mobile. | 25 | 1.00 | 6.00 | 4.4800 | 2.16256 |
| 3 I felt comfortable working with mobile | 25 | 1.00 | 6.00 | 5.3600 | 1.11355 |
| 4 Now that I have finished this test with mobile, I would feel at ease in taking other tests using mobiles. | 25 | 1.00 | 6.00 | 5.2400 | 1.23423 |
| 5 I felt uncomfortable using mobile for testing. | 25 | 1.00 | 6.00 | 1.9600 | 1.64520 |
| 6 I felt aggressive and hostile toward mobile when taking the test with it. | 25 | 1.00 | 6.00 | 1.6400 | 1.28712 |
| 7 I got a sinking feeling when I saw that I had to use mobiles for test | 25 | 1.00 | 6.00 | 2.2800 | 1.74452 |
| 8 Taking a test with a mobile made me nervous. | 25 | 1.00 | 6.00 | 1.9200 | 1.55242 |
| 9 I would have felt better if I had taken paper and pencil test instead of mobile based test. | 25 | 1.00 | 6.00 | 2.2800 | 1.76824 |
| 10 Using mobile made me uneasy and confused | 25 | 1.00 | 6.00 | 2.3200 | 1.72530 |
| Valid N (listwise) | 25 | | | | |

The items of the questionnaire were divided into 2 sections. First four items were based on positive statements specifying positive attitude of learners towards Mobile Based Tests and the items 5 to 10 were negative statements.

As it is shown in the table 4.6, the mean scores for the items stating positive sentences are consistently higher than the mean scores for the negative ones. Results in the above table showed that from the respondents' perspective, the most widespread status with the Mobile Based Test is the success it made in the item 4 stating that " I felt comfortable working with mobile " with mean of (5.36), followed by the items 1 saying "Using the mobile for taking a test did not scare me at all" and "Now that I have finished this test with mobile, I would feel at ease in taking other tests using mobiles" which has been specified to be the following utmost important item with mean 5.24, this means that MBT has attracted the majority of examinees' awareness and resulted in a positive attitude.

Qualitative Analysis

The qualitative section for answering the second question of the study was analyzed by employing narrative research design based on five open ended questions which were answered by the respondents based on whatever they felt and experienced throughout participating in 10 mobile based tests. The more frequently repeated answers to each question are provided in the table 4.8.

Table 8. Open-ended questions and answers

| Questions | Answers |
|--|--|
| What did you like most about testing by mobile? | faster speed, immediate feedback, attractive atmosphere of application, , not dealing with paper and pen, paper saving |
| What did you dislike most about testing by mobile? | Some difficulties in accessing the application, |
| Do you prefer taking the tests using mobiles or paper and pen? | Yes: 23 students No: 2 students |
| Why would you choose a mobile based test over paper and pencil based test? | easiness, immediate feedback, attractive atmosphere of application, more concentration, |
| Do you have any other comment on how you felt about testing by mobile? | Cheating was not possible with using mobile application, motivation to participate in mobile testing session was higher than paper based testing, preferring to take other tests using mobile phones |

The answers to these open ended questions revealed that participants of the FMA group preferred mobile based assessment over paper based one. As it is shown in the table 4.7, easiness, faster speed, attractive atmosphere of application, and immediate feedback provided by application were all the advantages of mobile based assessment from the view point of participants. Some students also asserted that they were more motivated to participate in mobile based tests compared to paper based ones. On the other hand, the only drawback of using mobile for testing was that students had difficulty in accessing the application; consequently, it caused some problems for test takers to answer the questions in the required time. However, the mentioned problem can easily be solved with provision of a wireless network accessible to all students. All in all, considering both qualitative and quantitative data received from the participants in FMA group, it can be concluded that they had a positive attitude towards mobile based form of assessing vocabulary.

DISCUSSION

Technology, nowadays, is considered as one main element of all English classes and has influenced the process of language learning and teaching. One section of English classes where technology is proposed to be used is assessment. The application of computers in assessment has been studied before, but nowadays computers have been replaced by mobiles which are used in our everyday lives. As mentioned by Klimova (2017), currently, there is an increasing trend in the shift from the use of traditional technologies such as a desktop computer towards the use of mobile technologies such as mobile phones or smartphones. Therefore, the application of such new technologies in English classes, especially the assessment part, is proposed throughout this study.

One important section of each class is assessment that helps both students and teachers enjoy a beneficial class. Appropriate assessment helps students meet certain standards and be aware of the main purpose of their lessons. On the other hand, teachers use the assessment data to identify strengths and weaknesses in student performance, and to improve the quality of teaching and learning. Today, with the rise of post method pedagogy ideologies, formative assessment has replaced summative assessment and has been considered by more teachers (Black & Wiliam, 1998). In other words, assessment of learning (summative assessment) has shifted to assessment for learning (formative assessment), where assessment is in support of learning.

As it was mentioned by Khodabakhshzadeh, Kafi & Hosseinnia (2017), formative assessment is a new concept in Iran and most of the teachers have not enough literacy of this kind of assessment. The majority of teachers still prefer the paper and pencil tests and summative ones. Therefore, to help teachers apply formative assessment which has been proved to be a substitution for summative assessment in English classes and to help teachers overcome the difficulties of administrating formative assessment, a simpler method of formative assessment based on technology was proposed in this study.

To examine the effect of employing mobile based formative assessment in English classes and especially in vocabulary gain of EFL learners, two groups of students which were assigned to two different classes before the study were selected and the vocabulary skill of each group was assessed in different ways. The results of this process of applying these types of assessments answered the research questions of the present study.

Briefly, the results of the present study exhibited the positive affect of applying mobile in assessing vocabulary skill of pre-intermediate English learners. As it was shown, formative assessment based on mobile applications was significantly more effective in vocabulary gain of pre-intermediate EFL learners than formative paper based assessment. These results are in line with most previous studies, some of which will be mentioned in what follows.

The application of mobile phones in EFL classrooms was investigated before by many researchers like Liu and Wu (2016) whose findings indicated that the LINE-based learning activities helped students familiarize themselves with English vocabulary and improve their English reading comprehension. Throughout this study, the researcher moved further and examined the role of such technology in assessment part of EFL classes and it was concluded that assessment can also be significantly influenced by application of mobile technology.

The results of the present study is also in accordance with Rakhyoot and Weir's (2014) who empirically investigated the effectiveness of on-line formative assessment (OFA) and feedback in the context of an EFL

course for postgraduate students at the University of Strathclyde. They concluded that on-line formative assessment with in-question feedback affords a worthy contribution to improving students' learning.

This study also supported the results obtained from Tarighat and Khodabakhsh's study suggesting that MALA (Mobile Assisted Language Assessment) be used alongside other forms of assessment to form part of the learners' final score just like Zheng, Su & Lian's (2014), who developed an Online Formative Assessment System for a Chinese EFL course to support conducting formative assessment activities and further improving learners' English language skills.

On the other hand, while Gordon (2015) investigated the possibility of a difference between actual performance in Mobile Device Testing (MDT) versus Paper Based Testing (PBT), the present study tried to investigate the same possibility where both modes of assessment were conducted in a formative way and it was concluded that formative assessment based on mobile phones can improve vocabulary learning of students significantly. Gordon, however, reported that although testing based on mobile phones could be regarded as an alternative for paper based testing, no significant difference between them was witnessed. Therefore, contrary to the Gordon's study, the present study showed that using mobile phones in assessment can lead to a significantly better learning of vocabulary. One main reason leading to such difference between the results of this study and Gordon's may be is the nature of formative assessment in being equipped with immediate feedback provided after each test.

In fact, one of the factors leading to better vocabulary learning in FMA group could be the immediate feedback which was provided by the Socratic application after each test. The students, then, became aware of their vocabulary knowledge and tried to study more for next test while in the other group, formative assessment was applied based on paper and pen and the students were not provided with immediate feedback. Actually, assessment with the provision of feedback (as its main distinctive feature) has an enormous advantage since students are informed about the continuous development in their achievement. Ainsworth (2006) alludes to the dissimilarity between tests and assessments by arguing that the latter can: "a) motivate students to be more engaged in learning, b) help students develop positive attitudes toward a subject, and c) give students feedback about what they know and can do" (p.22). In other words, the provision of feedback for both students and teachers during the term is necessary and leads to a better learning and teaching.

Although the FPA group was also provided by feedback in this study, it took a long time for instructor to correct the papers after the test. Therefore, time distance between each test and the feedback provision in FPA group was not short enough to motivate learners to study more after each test.

As stated by Kilickaya (2016), based on previous research on providing immediate feedback (Jonson, 2006; Chappuis & Chappuis, 2010; Pennebaker, Grosling & Ferrell, 2013; Fulcher & Owen, 2016), learners can benefit more from feedback on strengths and weakness, and in this perspective, the application was found to be highly efficient. Most assessments are done summatively (Mumm, Karm & Remmik, 2016). However, as formative assessment is believed to be providing more detailed feedback on the learners' progress when compared with the summative approach, Socratic application, based on the participants' views and experience in the classroom, seems to help teachers conduct formative assessment to determine students' weaknesses and strengths by automated scores, which was in accord with Fageeh's study (2015) and Rakhyoot and Weir's (2014).

The positive attitude of learners towards mobile based assessment would be considered as one main factor leading to such significant difference between FMA group and FPA one. The positive attitude of learners was concluded by analyzing the results of an attitude questionnaire distributed among the participants of the group whose vocabulary gain was assessed formatively based on a mobile application specified for testing. The results of the third question of this study are supported by the study done by Tarighat and Khodabakhsh (2016) who based on the interviews about the learners' attitude towards mobile based assessment for speaking declared that Mobile Assisted Language Assessment can be used alongside other forms of assessment to form part of the learners' final score. The only difference was that the present study dealt with assessing vocabulary gain of learners. Therefore, beside mobile based assessment of speaking, EFL learners have positive attitude towards vocabulary mobile based assessment too.

The other important factor helping FMA students learn better was that they were assessed with a device that was attractive to them and they also dealt with that in their everyday lives and of course they had a good attitude toward applying them in education. That is why they felt more comfortable using these devices and therefore they performed better and with higher level of energy. The students did not also need to take the mobile based tests in an especial laboratory. This could be regarded as an important advantage of mobile based assessment over computer based testing. Many other factors like time saving, atmosphere of testing sessions, and the attractions of the Socrative mobile application were beneficial for students using mobile based assessment to learn better.

CONCLUSION

Assessment as one main section of EFL classes was proved to be influenced by computers. CBT which stands for computer based test, according to previous studies, can replace paper based test appropriately. Thus, through this study the application of another technological device named mobiles in assessment was examined. Not necessary to say that because of the importance of vocabulary learning and also to narrow down the domain of study to get more reliable results, the vocabulary subskill was assessed in this study.

Based on the results of this study, participants who were assessed formatively using mobile and an application designed for testing named Socrative, were significantly better vocabulary learners after ten sessions of treatment than other group in which vocabulary gain of learners was assessed formatively based on paper and pen. Therefore, it was concluded that applying mobile devices and testing applications in assessment section of EFL classes in a formative way, improves vocabulary leaning of participants better than those who are assessed formatively based on paper and pen.

Formative assessment which is defined as the process of assessing students during the term for better learning is ignored nowadays by English teachers. Applying technology to facilitate this type of assessment in EFL class was suggested in this study. Mobile phone as a device which is considered as an inseparable part of human life was used as a means of performing a better vocabulary formative assessment. Immediate feedback as an advantage of mobile based formative assessment beside attractive atmosphere of mobile testing application and faster process of test taking differentiated this type of assessment with paper based ones where it took longer time for teacher to correct the papers and give the feedback to students. Being aware of their results in one test immediately, students tried more for next tests. Therefore, they learned more and performed better in consequent tests and posttest.

Regarding implications and with considering the conclusion of the study using mobile applications which are designed and produced to facilitate the process of testing is implicated. But more clearly, the main implication of this study is the application of formative assessment based on mobile application which can give immediate feedback, put learners in an attractive situation better than conventional and boring paper based testing sessions, and help teachers save their time and design a better plan for the future of their classes based on their received results. University students studying general English participated in this study. English institutions and schools can also apply this type of assessment and encourage their students use their mobiles in a beneficial way instead of forbidding them from carrying their mobiles.

Like any other research, this study also deals with some limitations and delimitations. First, based on the research questions of this study, 40 students from both genders took part in this study and their ages ranged from 18 to 22. These participants were from two intact classes studying general English at Chabahar Maritime University, and hence subject selection was not random. On the other hand, the research was intentionally delimited to EFL learners at the pre-intermediate language proficiency level to neutralize the effect of their language proficiency. In order to make the results more generalizable, other studies should include the participants of different language proficiency levels. Among all mobile apps which are used for testing, only one, Socrative App, was applied in the experimental class. Moreover, all tests in all groups were delimited to vocabulary and other areas of language were not included in this study.

BIODATA and CONTACT ADDRESSES of AUTHORS



Dr. Nahid YARAHMADZEHI, is a Professor of Linguistics at Management and Humanities Faculty, English Language Department, Chabahar Maritime University. Dr. Yarahmadzahi gained her Ph.D. in Linguistics at March, 2010. Her academic interest areas are theoretical, social, as well as Educational Linguistics, psycholinguistics and bilingualism, e-learning, and use of internet in education. She has over than 7 journal articles published in international indexes, and other national and international articles, papers submitted to international meetings.

Dr. Nahid YARAHMADZEHI

English Language Department, Management and Humanities Faculty

Address: Chabahar Maritime University, Postal Code: 9971756949, Chabahar, Iran

Phone: +985435322753

E-mail: venayarahmadi@gmail.com



Mostafa GOODARZI, received his Master's degree in English Language Teaching at Management and Humanities Faculty, English Language Department, Chabahar Maritime University. Mr. Goodarzi gained his Master's degree in English Language Teaching at March, 2019. His academic interest areas are use of internet in education, online testing, MALL, English language teaching methodology and applied linguistics. He is an English instructor in high schools, universities and institutions.

Mostafa GOODARZI

English Language Department, Faculty of Management and Humanities

Address: Chabahar Maritime University, Postal Code: 9971756949 Chabahar, Iran

Phone: +98 9354703574,

E-mail: mostafa.gudarzi1993@gmail.com

REFERENCES

- Ainsworth, L. (2006). *Common formative assessments: How to connect standards based instruction and assessment*. Thousand Oaks, California: Crown Press.
- Al-Amri, S., 2007. Computer-based vs. paper-based testing: are they the same? IN: *Khandia, F. (ed.)*. 11th CAA International Computer Assisted Conference: Proceedings of the Conference on 10th & 11th July 2007 at Loughborough University, Loughborough, 3-13.
- Allagui, B. (2014). Writing through WhatsApp: an evaluation of students writing performance. *International Journal of Mobile Learning and Organization*, 8 (3/4), 216 – 231.
- Bachman, L. (2004). *Statistical analysis for language assessment*. Cambridge: Cambridge University Press.
- Baleghizadeh S., & Oladrostam, E. (2010). The effect of mobile assisted language learning (MALL) on grammatical accuracy of EFL students. *MEXTESOL Journal*, 34(2), 1-10
- Black, P. & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*. 5(1), 7 - 74
- Brown, G. A., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. New York: Routledge.
- Brown, S., Race, P. & Smith, B. (2000). *500 tips on assessment. (1st ed.)* London: Kogan Page Limited.
- Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge: Cambridge University Press.
- Chen, C. M., & Hsu, S. H. (2008). Personalized intelligent mobile learning system for supporting effective English learning. *Educational Technology & Society*, 11(3), 153-180.

- Chickering, A. W., & Ehrmann, S. C. (1996). *Implementing the seven principles: Technology as lever*. AAHE Bulletin, 49, 3-6
- Chinnery, G. (2006). Going to the MALL: mobile assisted language learning. *Language Learning and Technology*, 10(1), 9–16.
- Ducate, L., & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning & Technology*, 13(3), 66-86
- Duncan, A. & Dunn, W. (1988). *What Primary Teachers should know about Assessment*. London: Hodder & Stoughton.
- Fleming, S. & Hiple, D. (2004). Foreign language distance education at the University of Hawai'i. In C. A. Spreen, (Ed.), *New technologies and language learning: issues and options* (Tech. Rep. No.25) (pp. 13-54). Honolulu, HI: University of Hawai'i, Second Language Teaching & Curriculum Center.
- Khodabakhshzadeh, H., Kafi, Z., & Hosseinnia, M.,. Investigating EFL Teachers' Conceptions and Literacy of Formative Assessment: Constructing and Validating an Inventory, *International Journal of Instruction*, January 2018.
- Kukulka-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL*, 21(2), 157-165.
- Kukulka-Hulme, A., Norris, L. and Donohue, J. (2015). Mobile pedagogy for English language teaching: a guide for teachers. *British Council ELT Research Papers*, [pdf] 14.07. Available at: http://englishagenda.britishcouncil.org/sites/default/files/attachments/e485_mobile_pedagogy_for_elt_final_v2.pdf
- Khoshsima, H., & Hashemi Torujeni, M. (2017). Comparability of Computer-Based Testing and Paper-Based Testing: Testing Mode Effect, Testing Mode Order, Computer Attitudes and Testing Mode preference. *International Journal of Computer (IJC)*. 24(1), 80-99.
- Klimova, B. (2017). Mobile phones and/or smartphones and their apps for teaching English as a foreign language, Education and Information Technologies, *The Official Journal of the IFIP Technical Committee on Education*. doi: 10.1007/s10639-017-9655-5.
- Liu, I., Chen, M.C., Sun, Y.S., Wible, D., & Kuo, C. (2010). Extending the TAM model to explore the factors that affect intention to use an online learning community. *Computer & Education*, 52, 600-610. doi:10.1016/j.compedu.2009.09.009.
- Rezaei, M., & Golshan, M. (2015). Computer Adaptive Test (CAT): Advantages and Limitations. *International Journal of Educational Investigations*. 2(5): 128-137.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19(2), 12–25.
- Sahan, O., Coban, M., & Razi, S. (2016). Students learn English idioms through WhatsApp: Extensive use of smartphones. *Erzincan University Journal of Education Faculty*, 18(2), 1230-1251.
- Sorayyaei Azara, A., & Nasiri, H. (2014), Students learn English idioms through WhatsApp: Extensive use of smartphones. *Uremia University, Iran*. doi: 10.1016/j.sbspro.2014.03. Published by Elsevier Ltd.
- Spector, J. M., Ifenthaler, D., Samspon, D., Yang, L., Mukama, E., Warusavitarana, A., et al. (2016). Technology enhanced formative assessment for 21st century learning. *Educational Technology & Society*, 19(3), 58_71.
- Stobart, G. & Gipps, C. (1997). *Assessment: A teacher's guide to the issues*. London: Hodder & Stoughton
- Stahl, S.A. (2005). *Four problems with teaching word meanings (and what to do to make vocabulary an integral part of instruction)*. Mahwah, NJ: Erlbaum.
- Tseng, W. (2017). Measuring English vocabulary size via computerized adaptive Testing. *Computers & Education*, 97, 69-85.

- Trumbull, E. & Lash, A. (2013). *Understanding formative assessment: Insights from learning theory and measurement theory*. San Francisco, CA: WestEd.
- Taki, S. & Khazaei, S. (2011). Learning Vocabulary via Mobile Phone: Persian EFL Learners in Focus. *Journal of Language Teaching and Research*, 2(6), 1252-1258.
- Varier, D. (2015). *A Case Study Examining Formative Assessment in a Postsecondary English Language Program*, Virginia Commonwealth University, Richmond, Virginia, Department of Foundations of Education.
- Yang, J. (2013). Mobile assisted language learning: review of the recent applications of emerging mobile technologies. *English Language Teaching*, 6(7), 19-25.