

## Quizzes in Every Other Session Improve Undergraduate EFL Learners' Pronunciation Achievement

Mohammad Reza Ghorbani\*

University of Bojnord, Iran

Corresponding Author: Mohammad Reza Ghorbani, E-mail: mrg872@gmail.com

### ARTICLE INFO

#### Article history

Received: June 24, 2017

Accepted: September 18, 2017

Published: October 31, 2017

Volume: 8 Issue: 5

Advance access: October 2017

Conflicts of interest: Non

Funding: None

#### Key words:

Frequent Quizzes,

EFL Teaching,

EFL Learning,

EFL Testing,

Pronunciation Achievement

### ABSTRACT

The proponents of frequent quizzes claim that they stimulate students and have a positive effect on their learning, while the opponents argue that too frequent quizzes might frustrate students and hinder their learning. This study examined the effect of frequent quizzes on Iranian undergraduate English as a Foreign Language (EFL) learners' pronunciation achievement. The nonequivalent group, pretest-posttest design was employed to study two classes of English literature and English teaching students, who were taking the Phonology Course, at Kosar University of Bojnord (KUB) as the experimental group (EG) and control group (CG) respectively. Two 40-item pronunciation tests were developed based on the 3<sup>rd</sup> edition of the book *Ship or Sheep* written by Baker (2006). The reliability of the tests was estimated 0.78 and 0.81 respectively through KR-21 formula. After the pretest administration, both groups were exposed to the same activities; however, only the EG took the quizzes every other session. At the end of the training program, the pretests were rearranged and used as the posttests. The results of the independent samples t-tests from the posttests revealed that the EG had a better performance than the CG suggesting that EFL learners' pronunciation achievement can improve if quizzes are used every other session.

### INTRODUCTION

Applying methods to improve learning has always been desirable for teachers and students. Evaluation as part of the learning process can be used to monitor learning and teaching processes. As pointed out by Pourahmad, Fadaei, Tadayon, and Zabetian (2013), students usually have to study a large volume of materials to pass the final exams. Although they have enough time during the course, they begin to study the materials taught during the course just a few days before the final exams. They believe that students are less anxious and stressed about the exam during the course and they can use this time for effective revision of the subjects. They state that if the content is studied during the course, it will be more stable and lasting.

According to Adkins and Linville (2017), the effect of testing frequency has already been studied; however, no definite consensus has been formed on the number of quizzes in a course and its impact on student performance. They argue that it is difficult for instructors to conduct frequent testing; therefore, they probably administer only midterm and final exams in their courses. Lack of enough faculty resources (Kuo & Simon, 2009) and consumption of valuable instruction time (Mines, 2014) are mentioned in the literature as the main reasons for less frequent quizzes in the classroom.

Studies on frequent testing have found conflicting results. Some have found that it positively influence student performance (Gholami & Morady Moghaddam, 2013) while others have found that less frequent testing does not make a difference in student learning (Mines, 2014). This study was an attempt to examine the impact of quizzes in every other session on improving undergraduate EFL learners' pronunciation achievement in a phonology course taken by undergraduate EFL students in the Iranian context.

Frequent quizzes are one way of encouraging students to study what has been taught in class during the previous session. The term frequent testing has been defined as a weekly examination (Keys, 1934), a daily assessment (Dineen, Taylor, & Stephens 1989), a monthly test (Kling, Miller, & Reardon, 2005). According to Mousavi (2009, p. 546), a quiz is "the most common classroom test which is a compromise between short-term, subjective evaluation based on daily work and the longer term achievement test". He believes that quizzes have a positive effect on students' learning, reduce test anxiety, serve as a review at the beginning of a class, raise students' attention at the end of a class, give feedback to teachers, encourage students to regularly study their language, enable teachers to acquaint their students with sample items used in formal tests, and help students diagnose their limitations.

As pointed out by Gholami and Morady Moghaddam (2013, p. 36), "it is generally assumed that quizzes are useful tools to enhance learning and consolidate what has been taught". Brown (2004, p. 29) considers quizzes as "the information that washes back to students in the form of useful diagnoses of strengths and weaknesses". The effect of testing on teaching and learning, which can be either beneficial or harmful, has been referred to as backwash in general education and washback in the field of language teaching (Ghorbani & Neissari, 2015; Ghorbani, 2012).

Geist and Soehner (1997) reported the significant and positive effect of frequent quizzes on student performance. Their study showed that increasing the number of quizzes led to the students' better performance. Roediger and Karpicke (2006) studied the effect of frequent testing on undergraduate university students' retention of information and found that the frequently tested students during the course had a better performance. Kamuche (2005) found that weekly tests improve students' performance.

Basol and Johanson (2009) used 78 studies to investigate the effect of different testing frequencies on student achievement based on a meta-analysis. They classified the studies into high, medium, and low frequency. They found that frequent testing was beneficial regardless of their high or low frequency. Phelps (2011) also carried out a meta-analysis, based on the studies conducted until 2009, on the effect of testing on achievement and found that testing has the strongest positive effect on achievement if it is followed by feedback. The result of his study indicated that the effect of testing on achievement ranges from moderate to strong and high stakes tests strongly affect achievement.

Frequent quizzes can motivate students and increase their classroom attendance (Zarei, 2008). In addition, they prepare students for high-stakes tests and boost their long-term retention of the material (Johnsom & Kiviniemi, 2009). They help teachers make sure that their students are doing their assignments (Weinstein & Wu, 2009). However, according to Gholami and Morady Moghaddam (2013), the process of test administration and scoring is time-consuming and frequent testing leads to time constraint for instruction. Furthermore, Marshall (2007) argues that testing a lot does not lead to useful learning in the long-term. He thinks that since teachers teach and students read to the test, learning will be restricted to what appears on the test. For example, in the Iranian context, due to the washback effect of written exams, little attention is given to teaching pronunciation in the public education system (Ghorbani & Neissari, 2015; Ghorbani, 2012; Ghorbani, 2011; Hayati, 2010; Hosseini, 2007).

Gholami and Moradi Moghaddam (2013) studied the effect of weekly quizzes and found that they are more effective than the mid-term exam only. Mirhassani and Rahimpour (2003) found that frequent quizzes improve Iranian EFL learners' performance on summative achievement tests. Ballard and Johnson (2004) also found that frequent quizzes influence learning performance. Tuckman (2008) indicated that frequent testing motivates students and learners who are frequently tested outperform those who are not on examinations.

Shirvani (2009) investigated the effect of daily quizzes on students' final exam scores and homework grades. He found that daily quizzes had a significant effect on both of them. Momeni and Barinani (2012) investigated the effect of weekly and bi-weekly quizzes on Iranian pre-intermediate EFL learners' language achievement. They found that the group who took quizzes outperformed those who did not. They also found that the group with weekly quizzes had a better performance than the group with bi-weekly quizzes on the final exam

However, quizzes may not always be useful. In contrary to the above-mentioned studies, other researchers found contradictory results. For instance, Zraggen (2009) investigated the effect of weekly versus bi-weekly testing on students' learning and retention. The group who received the tests on a bi-weekly basis outperformed the group who took the tests every week on both the final exam and the retention test one month later. Furthermore, the results of a study by Pourahmad, et al. (2013) showed that quizzes during the course of study have no effect on the final score of the students. Zamini, Khademerfan, Rahmani, Khodavaisy, and Davari (2013) also found that taking frequent quizzes was not associated with higher final scores in comparison to the regular training technique. Therefore, more research in this area is warranted.

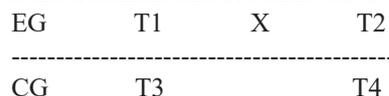
To investigate the effect of frequent quizzes on undergraduate EFL learners' pronunciation achievement, the following research questions and null hypotheses were formulated:

1. Do frequent vowels quizzes increase undergraduate EFL learners' pronunciation achievement?
2. Do frequent consonants quizzes increase undergraduate EFL learners' pronunciation achievement?
1. Ho: Frequent vowels quizzes do not increase undergraduate EFL learners' pronunciation achievement.
2. Ho: Frequent consonants quizzes do not increase undergraduate EFL learners' pronunciation achievement.

## METHOD

### Research Design and Subjects

The participants were 59 Iranian undergraduate EFL learners at KUB. They were all female students who attended a two-credit pronunciation course and received one and a half hours of instruction a week. Since random assignment was not possible, the nonequivalent group, pretest-posttest design was employed in this study. That is, subjects were tested in existing groups. The following diagram summarizes this quasi-experimental design in which the dotted line represents non-equivalent groups. Both groups were measured before and after the treatment. Only one group received the treatment. In this diagram EG and CG stand for experimental and control groups respectively. T1 and T3 stand for the tests before applying the treatment. T2 and T4 stand for the tests after the treatment and X stands for treatment.



## Materials and Instruments

The material was based on the 3<sup>rd</sup> edition of the book *Ship or Ship* written by Baker (2006) including a set of four audio CDs which give lots of listening and pronunciation practice. The researcher used the first two audio CDs related to 22 units in the first section of the book under the title of Vowels and the second two audio CDs related to 28 units in the second section of the book under the title of Consonants. According to the author, this fully-revised and updated edition provides systematic practice of English pronunciation, with an emphasis on minimal pairs, through a wide variety of interesting exercises and activities. The new edition has been re-written to make it suitable for either self-study or classroom use. It trains students to recognize and produce English sounds by helping them make the distinction between similar sounds. Its stand-alone units allow learners to focus on sounds that they find difficult. Each unit offers comprehensive practice of sounds, with additional work on stress and intonation. The book recommends students to visit [www.cambridge.org/elt/shiporsheep](http://www.cambridge.org/elt/shiporsheep) for extra practice and web support.

Two 40-item tests were developed by the researcher based on word transcription and stress, listening to the dialogue and paying attention to the target sound as well as minimal pair exercises. The first test was based on the first section of the book under the title of Vowels and the second test was based on the second section of the book under the title of Consonants. The two tests were used both as the pre-test and the post-test. For the pilot test, a group of university students similar to those of this study responded to the items and helped the researcher establish the reliability of the tests which were estimated 0.78 and 0.81 respectively through KR-21 formula. Each test included the following subsections:

1. Transcribe the following words using phonetic symbols and put the stress mark on the correct syllables. (12 points)
2. Put the following words under the right sound. (8 points)
3. Listen and write down what you hear. (20 points).

## Procedure

In this study, there were two EFL classes with 33 learners in one class and 26 learners in the other. After administering the first pre-test, the 8-week treatment on English vowel sounds began. After administering the first post-test and the second pre-test, the 8-week treatment on English consonant sounds began. The whole process of the study began from February 7 to June 2, 2015. The subjects took part in one and a half hour classes one day a week in the afternoon during which all the English vowels and consonants were introduced to them. The whole class time was allocated to teaching English vowels and consonants.

Both classes were exposed to authentic pronunciation from the beginning of the course. In both classes, the subjects listened to the same audio CDs while looking at their books and were provided with explicit information on pronunciation if needed. The subjects were allowed to look

certain words up and check their phonemic transcription in their dictionary. Then, they were asked to read the exercises aloud with correct pronunciation after five minutes of group work, and the researcher would try to help them if necessary. The only difference between the two groups was that the EG took four quizzes (one quiz every other week) during the treatment while the CG had no quizzes. The CG spent all the class time on listening and pronunciation practice.

At the end of each treatment, the 40-item pre-tests were used again as the post-tests to see if there was any significant difference between the two groups' performances. The pre-tests and post-tests were identical but the arrangement of the items was different in the post-tests. Since there was an interval of two months between the two tests, the post-tests were less likely to be influenced by the subjects' memory. The subjects listened to each item three times and wrote the correct answer down during the pre-tests and post-tests. Each correct answer received one point.

## STATISTICAL ANALYSIS AND RESULTS

To explore the effect of the treatment on the EFL learners' academic achievement, the data were subjected to statistical analysis. To answer the first research question (Do frequent vowels quizzes increase undergraduate EFL learners' pronunciation achievement?), after rating the subjects' performance, the raw scores taken from the pre-test and post-test were submitted to the computer software Statistical Package of Social Sciences (SPSS version 22), using a t-test. Independent samples t-test was conducted to compare the possible differences between the means of the two groups based on the gain scores from the post-test. The calculation indicated that frequent vowels quizzes were more effective. The following tables indicate the summary of the t-tests.

An independent samples t-test was conducted to compare the scores of the two groups before the treatment. First, the Levene's Test for Equality of Variances was checked. If the Levene's Test is significant ( $p < .05$ ), the two variances are significantly different. If it is not significant ( $p > .05$ ), the two variances are approximately equal. In this case, since the Levene's test was not significant ( $p = 0.43 > 0.05$ ), it was assumed that the variances were approximately equal. Next, the results of the t-test were checked. If the variances are approximately equal, the top line is read. If the variances are not equal, the bottom line is read. Based on the results of the Levene's test, it was known that the two groups had approximately equal variances on the dependent variable, so the top line was read.

As indicated in Table 1, there was no significant difference between the EG ( $M = 10.48$ ,  $SD = 1.75$ ) and the CG [ $M = 11.19$ ,  $SD = 1.57$ ;  $t(57) = -0.78$ ,  $p > .05$ ] before the treatment.

The second independent samples t-test was conducted to compare the scores of the two groups after the treatment (the first post-test). First, the Levene's Test for Equality of Variances was checked. Since the Levene's test was not significant ( $p = 0.87 > .05$ ), it was assumed that the variances were equal. Next, the results of the t-test were checked. Based on the results of the Levene's test, it was known that the two

groups had equal variances on the dependent variable, so the top line was read.

As indicated in Table 2, there is a significant difference between the gain scores for the EG ( $M = 31.81, SD = 2.96$ ) and the gain scores for the CG [ $M = 28.88, SD = 3.65; t(57) = 3.65, p < .05$ ]. This final result shows that the mean score of the EG after the treatment is more than the CG. Since there is a significant difference between the means of the two groups, the first null hypothesis (Frequent vowels quizzes do not increase undergraduate EFL learners' pronunciation achievement.) is rejected. Therefore, the effectiveness of frequent vowels quizzes is supported.

To answer the second research question (Do frequent consonants quizzes increase undergraduate EFL learners' pronunciation achievement?), after rating the subjects' performance, the raw scores taken from the second pre-test and post-test were submitted to the computer software SPSS (version 22), using a t-test. Independent samples t-test was conducted to compare the possible differences between the means of the two groups based on the gain scores from the second post-test. The calculation indicated that frequent consonants quizzes were more effective. The following tables indicate the summary of the t-tests.

An independent samples t-test was conducted to compare the scores of the two groups before the treatment. First, the Levene's Test for Equality of Variances was checked. If the Levene's Test is significant ( $p < .05$ ), the two variances are significantly different. If it is not significant ( $p > .05$ ), the two variances are approximately equal. In this case, since the Levene's test was not significant ( $p = 0.78 > 0.05$ ), it was assumed that the variances were approximately equal. Next, the results of the t-test were checked. If the variances are approximately equal, the top line is read. If the variances are not equal, the bottom line is read. Based on the results of the Levene's test, it was known that the two groups had approximately equal variances on the dependent variable, so the top line was read.

As indicated in Table 3, there was no significant difference between the EG ( $M = 11.21, SD = 1.43$ ) and the CG [ $M = 11.42, SD = 1.33; t(57) = -0.57, p > .05$ ] before the treatment.

**Table 1.** The independent samples t-test for the experimental and control groups (the first pre-test)

Group	N	Mean	Standard deviation	df	t	Sig.
Experimental	33	10.84	1.75	57	-0.78	0.43
Control	26	11.19	1.57			

**Table 2.** The independent samples t-test for the experimental and control groups (the first post-test)

Group	N	Mean	Standard deviation	df	t	Sig.
Experimental	33	31.81	2.96	57	3.65	0.001*
Control	26	28.88	3.19			

\*Sig.  $p < 0.05$

The second independent samples t-test was conducted to compare the scores of the two groups after the treatment (the second post-test). First, the Levene's Test for Equality of Variances was checked. Since the Levene's test was not significant ( $p = 0.73 > .05$ ), it was assumed that the variances were equal. Next, the results of the t-test were checked. Based on the results of the Levene's test, it was known that the two groups had equal variances on the dependent variable, so the top line was read.

As indicated in Table 4, there is a significant difference between the gain scores for the EG ( $M = 32.18, SD = 2.76$ ) and the gain scores for the CG [ $M = 29.11, SD = 3.14; t(57) = 3.98, p < .05$ ]. This final result shows that the mean score of the EG after the treatment is more than the CG. Since there is a significant difference between the means of the two groups, the second null hypothesis (Frequent consonants quizzes do not increase undergraduate EFL learners' pronunciation achievement.) is rejected. Therefore, the effectiveness of frequent consonants quizzes is supported.

## CONCLUSION

This study was an attempt to investigate the effect of frequent quizzes on Iranian undergraduate EFL learners' pronunciation achievement. The findings indicated that the performance of the group who took quizzes every other session was significantly better than that of the control group. The results are in line with those of Geist and Soehren (1997) and Ballard and Johnson (2004) who found weekly quizzes enhance students' performance. It is also consistent with the findings of Kamuche (2005) in which the students who received weekly quizzes outperformed those who did not. It is also in keeping with the results reported by Gholami and Moradi Moghaddam (2013), Momeni and Barinani (2012), Shirvani (2009), Tuckman (2008), Roediger and Karpicke (2006), and Mirhassani and Rahimipour (2003).

There seems to be many factors and reasons that contribute to the positive effect of frequent testing on students' pronunciation achievement. Class attendance is supposed to be one of the reasons behind the success of frequent quizzes. Another reason is that quizzes provide students with a lot

**Table 3.** The independent samples t-test for the experimental and control groups (the second pre-test)

Group	N	Mean	Standard deviation	df	t	Sig.
Experimental	33	11.21	1.43	57	-0.57	0.56
Control	26	11.42	1.33			

**Table 4.** The independent samples t-test for the experimental and control groups (the second post-test)

Group	N	Mean	Standard deviation	df	t	Sig.
Experimental	33	32.18	2.76	57	3.98	0.000*
Control	26	29.11	3.14			

\*Sig.  $p < 0.05$

of extrinsic motivation and, therefore, increase their course grades (Zarei, 2008). Furthermore, as pointed out by Gholami and Moradi Moghaddam (2013), regular administration of quizzes to students will make them accustomed to the tests and reduce their anxiety before the final exam. This lower level of debilitating test anxiety, in turn, seems to have played a role in improving the learners' achievement.

Since the main use of quizzes in this study was to diagnose strengths and weaknesses of individual learners, they provided learners with valuable corrective feedback and helped them assess, evaluate, and improve their achievement. Quizzes served as a formative assessment by identifying the content areas that students had not learned well and merited further study. The researcher provided some elaboration on each item which enhanced students' learning.

However, the findings of this study are not consistent with those of Zraggen (2009) and Pourahmad, et al. (2013) who showed that quizzes during the course of study have no effect on the final score of the students. The results are not also in line with those of Zamini, Khademerfan, Rahmani, Khodavaisy, and Davari (2013) who found that taking frequent quizzes was not associated with higher final scores in comparison to the regular training technique. Furthermore, the findings are not in alignment with those of Mines (2014) who found no significant difference in learners' achievement with less frequent testing.

It seems that there is no conclusive evidence to determine whether frequent or infrequent quizzes are more helpful. This contradiction in results may be due to the students' general presumption that they have learned the content well during the course and they do not need any revision. Another reason is that students may not have taken quizzes seriously or may have allocated less time for the final exam preparation. Therefore, due to these contradictory results, further studies are recommended to confirm the effect of frequent quizzes. The results of this study may have implications for teachers, learners, material writers, test developers, and curriculum designers.

## REFERENCES

- Adkins, J. K. & Linville, D. R. (2017). Testing frequency in an introductory computer programming course. *Information Systems Education Journal*, 15(3), 22-28.
- Baker, A. (2006). *Ship or Sheep: An Intermediate Pronunciation Course*. Cambridge: Cambridge University Press.
- Ballard, C. L., & Johnson, M. F. (2004). Basic math skills and performance in an introductory economics class. *Journal of Economic Education*, 35(1), 3-24.
- Basol, G. & Johnson, G. (2009). Effectiveness of frequent testing over achievement: A meta-analysis study. *International Journal of Human Sciences*, 6(2), 99-120.
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. US: Longman.
- Dineen, P., Taylor, J., & Stephens, L. (1989). The effect of testing frequency upon the achievement of high school mathematics course. *School Science and Mathematics*, 89(3), 197-200.
- Geist, J. R. & Soehren, S. E. (1997). The effect of frequent quizzes on short- and long-term academic performance. *Journal of Dental Education*, 61(4), 339-345.
- Gholami & Morady Moghaddam (2013). The effect of weekly quizzes on students' final achievement score. *I. J. Modern Education and Computer Science*, 1, 36-41.
- Ghorbani, M. R. & Neissari, M. (2015). Washback effect of the Iranian concours on senior high school students' EFL learning activities. *The Iranian EFL Journal*, 5(1), 1-28.
- Ghorbani, M. R. (2012). *Washback effect of the university entrance examination: Iranian case*. Germany: Lambert Academic Publishing (LAP).
- Hayati, M. (2010). Notes on teaching English pronunciation to EFL learners: A case of Iranian high school students. *English Language Teaching*, 3(4), 121-126.
- Hosseini, S. M. H. (2007). ELT in higher education in Iran and India – A critical view. *Language in India*, 7, 1-11.
- Johnson, B. C. & Kiviniemi, M. T. (2009). The effect of on-line chapter quizzes on exam performance in an undergraduate Social psychology course. *Teach Psychology*, 36(1), 33-37.
- Kamuche, F. U. (2005). Do weekly quizzes improve student performance? *Academic Exchange Quarterly*, 9(3), 188-193.
- Keys, N. (1934). The influence on learning and retention of weekly as opposed to monthly tests. *Journal of Educational Psychology*, 25, 427-436.
- Kuo, T., & Simon, A. (2009). How many tests do we really need?. *College Teaching*, 57(3), 156-160.
- Kling, N., Miller, C., & Reardon, J. (2005). The impact of testing frequency on student performance in a marketing course. *Journal of Education for Business*, 81(2), 67-72.
- Marshall, B. (2007). A crisis for efficacy? *Education Review*, 20(1), 29-35.
- Mines, Jr., Richard O., "The Impact of Testing Frequency and Final Exams on Student Performance," *Proceedings of the ASEE Southeastern Section Annual Conference*, March 30-April 1, 2014; podium presentation by R. O. Mines, Jr., March 31, 2014, 10:20 to 10:40 am.
- Mines Jr, R. O. (2014). The impact of testing frequency and final exams on student performance. American Society for Engineering Education Southeast Section Conference. *Proceedings of the ASEE Southeastern Section Annual Conference*, March 30-April 1, 2014; podium presentation by R. O. Mines, Jr., March 31, 2014, 10:20 to 10:40 am.
- Mirhassani, A., & Rahimpour, S. (2003). The relationship between quiz, techniques, frequency of administration, and Iranian EFL learners' performance on summative achievement tests. *English Secretariat Quarterly Journal*, 10, 25-38.
- Momeni, A., & Barimani, S. (2012). The effect of testing frequency on Iranian pre-intermediate EFL learners' language achievement. *Journal of Academic and Applied Studies*, 2(10), 76-87.
- Mousavi, S.A. (2009). *A Dictionary of Language Testing*. Tehran: Rahnama Publication.
- Pourahmad M., Fadaei S., Tadayon S., & Zabetian H. (2013). The effect of quizzes on the students' scores in final exam. *Journal of Jahrom University of Medical Sciences*, 10(4), 37-41.

- Phelps, R. P. (2011). The effect of testing on student achievement. *International Journal of Testing*, 12(1), 21-43.
- Roediger, H. L. & Karpicke, J. D. (2006). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249-255.
- Shirvani, H. (2009). Examining an assessment strategy on high school mathematics achievement. *American Secondary Education*, 38(1), 34-45.
- Tuckman, B. W. (2008). Using frequent testing to increase students' motivation to achieve. *Journal of Experimental Education*, 66, 141-147.
- Weinstein, S. E. & Wu, S. W. (2009). Readiness assessment tests versus frequent quizzes: Student preferences. *International Journal of Teaching and Learning in Higher Education*. 21(2), 181-186.
- Zamini, G., Khademerfan, M. B., Rahmani, M. R., Khodavaisy, M. S., & Davari, B. (2013). Effects of frequent announced parasitology quizzes on the academic achievement. *Iranian Journal of Parasitol*, 8(4), 617-621.
- Zarei, A. A. (2008). On the Learnability of three categories of Idioms by Iranian EFL learners. *Journal of Humanities of the University of Kerman*, 2(2), 82-100.