

Effectiveness of Oral English Blended Teaching Supported by Digital Technology in Universities

Zhonger Wang^{1,2*}, Nurliyana Bukhari¹, Yang Han³, Qi Tang¹

¹School of Education and Modern Languages, Universiti Utara Malaysia, Sintok 06010, Malaysia

²School of Foreign Language, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China

³School of Languages, Civilisation and Philosophy, Universiti Utara Malaysia, Sintok 06010, Malaysia

*Corresponding author: Zhonger Wang, ZhongerWang@hotmail.com

Copyright: © 2023 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: An empirical study of oral English blended teaching was conducted in a university. The experiment class, which was under the blended teaching mode supported by digital technology, made significant improvements in language accuracy and range, discourse length and coherence, as well as flexibility and appropriateness. On the other hand, there was no significant improvement in the scores of the control class. Through the analysis of results by means of independent t-test and paired sample t-test, we drew the following conclusions: the oral performance of university students from the experiment class significantly improved; there was remarkable difference in the total oral English scores between the experiment class and the control class; the improvement in the total score and the score of each module showed that blended teaching had significant effect on the improvement of university students' oral English performance.

Keywords: Blended teaching; Digital technology; Oral English

Online publication: February 10, 2023

1. Introduction

The overall objective of this empirical study was based on the requirements of the College English curriculum syllabus in China. In light of the principle of interest, scientific nature, and communicativeness of the syllabus, this study aimed to help university students develop their oral English performance via a series of elaborate activities supported by digital technology in and after class. Oral English falls in the category of language output ^[1]. It is self-evident that the latter will be in place only if there is a large amount of language input ^[2]. Therefore, in this empirical study, the teaching of listening and speaking was conducted simultaneously; that is to say, the students learned both listening and speaking at the same time in the process of cooperative and interactive learning with the support of digital technology.

2. Research hypothesis

In order to verify the effectiveness of oral English blended teaching in universities, we conducted an empirical study over a period of one semester. The hypothesis in the study was as follows: compared with the traditional style, the blended one, which was supported up by digital technology, had more advantages, as we intended to achieve the goal of enhancing university students' oral English performance.

3. Experiment conditions

3.1. Subjects

We selected sophomores as our subjects. The reason was that they had relatively more class hours in English and plenty of extracurricular time, which ensured the smooth running of blended teaching in- and out-of-classroom settings.

For the experiment to be conducted smoothly, the class with high ownership rate of digital devices was selected as the experiment class. Of the 50 students in the experiment class, 90% of them had laptops equipped with wireless internet access, and all of them had mobile phones, which enabled them to play and record both audio and video clips. These students were post-00s and had an innate familiarity with digital learning.

As far as the randomly selected control class was concerned, its number and sex ratio were the same as those in the experiment class. Lectures were given to both the classes by the same lecturer. The class hours and teaching conditions were also same. Before the experiment, an English proficiency test in listening and speaking was undertaken. The scores were analyzed by means of independent sample t-test. From the statistical results, there was no significant difference in the scores between the two classes. In other words, both groups were basically at the same level and comparable.

3.2. Technical environment

Following the selection of the experiment class, it was found by survey that the majority of them had more free time at noon or in the evening. Data from a similar study conducted previously have also shown that both time intervals were the peak time for university students to surf the internet. Therefore, the guidance for students' extracurricular learning was mainly conducted during these two time intervals, *i.e.*, the students in the experiment class used mobile phones, laptops, i-pads, and other mobile devices for extracurricular learning under the guidance of their lecturer. While teaching, the lecturer employed both summative and formative evaluations to assess the students' learning effect.

3.3. Learning content and resources

The textbook used for both the experiment class and the control class was *Listening and Speaking for College English Courses* published by Shanghai Foreign Language Teaching Publishing Group. As for supplementary materials, original English films, news in Voice of America (VOA) Special English, BBC News, and newsreels, documentary shots, as well as animations were selected after certain factors were taken into consideration, such as the English proficiency and preferences of students.

The teacher provided teaching and learning resources through online learning platforms, such as "Super Star" and "U Campus." The adaptive content was accessible via the computer and mobile phone.

3.4. Evaluation tools

The oral test utilized in this empirical research covered a broad scope of topics and types. It was elaborately designed to evaluate the students' oral proficiency from different facets, especially their communicative competence. Unified scoring standards and methods were adopted. In order to maximize the internal consistency and reliability of the test, qualified examiners were selected.

3.4.1. Test content

In order to effectively assess communicative competence and the use of strategies in this regard, we divided the test content into three parts.

The first part was a 5-minute interview between the candidate and the examiner in the form of questions and answers. Questions relating to the life of the respective candidate, including their personal information,

hobbies, specialties, *etc.*, were posed by the examiner. The candidates were required to respond to these questions accordingly. This part mainly examined the respective candidate's fluency of oral expression and accuracy of pronunciation and intonation, thus reflecting the functionality of language.

The second part included individual speeches and group discussion for 1.5 and 4.5 minutes, respectively. The second part took about 10 minutes in total. This part assessed the respective candidate's ability in language organization and oral expression as well as whether the content was relevant and substantial.

In the third part, questions were asked again to determine the respective candidate's oral proficiency, especially in the aspect of communicative competence. The third part took about 5 minutes, in which the examiner played different roles to provide a communicative situation so that the candidate could communicate with the examiner, while playing certain appointed roles. This part examined the appropriateness of their oral expression and reflected the social and cultural characteristics of language.

3.4.2. Scoring standards and methods

The scoring standards of the oral test undertaken in this research fully reflected the priority of communicative competence. Importance was attached not only to the accuracy and informativeness in oral expression but also to the fluency and appropriateness of the spoken English. The scoring standards included three groups of items.

(1) Group 1: Language accuracy and range

- (a) Language accuracy refers to the precision of a candidate's pronunciation, intonation, grammar, and vocabulary ^[3].
- (b) Language range refers to the complexity and scope of vocabulary and grammatical structures used by the candidate ^[4].

(2) Group 2: Discourse length and coherence

- (a) The length of discourse refers to how much a candidate has contributed to the communication in the whole exam and how much he/she has spoken.
- (b) Coherence refers to the ability of a candidate to speak for a long time in coherent speeches.

(3) Group 3: Flexibility and appropriateness

- (a) Flexibility refers to the ability of a candidate to cope with different situations and topics.
- (b) Appropriateness refers to the ability of a candidate to use accurate and appropriate language in different situations ^[5].

The test was scored by analytical methods based on the three groups of items mentioned above. Each group had a full score of 5 points and was presented in a scale of five grades. The examiner scored each group according to the marking standards, in which higher scores meant better the performance. The total score was the sum of the scores for each group, and the final score of the students' oral performance was the average of the total scores by the three examiners.

3.5. Statistical analysis

We measured and analyzed the results in a quantitative approach. SPSS 22.0 was used to evaluate the test results, and two analysis methods were used to draw conclusions.

3.5.1. Paired sample t-test

Paired samples refer to two sets of data obtained from two tests on the same sample or from tests on two identical samples under different conditions. The paired sample test is carried out to infer whether the mean values of such two groups of data are significantly different ^[6].

The pretest and posttest scores of the experiment class were two sets of data obtained from two tests on the same sample. Therefore, paired sample t-test was used to evaluate whether the oral English proficiency scores in the pretest and posttest were significantly different [7].

Likewise, the pretest and posttest scores of the control class were two sets of data obtained from two oral tests on the same sample. Therefore, paired sample t-test was used to evaluate whether the oral English proficiency scores in the pretest and posttest were significantly different.

3.5.2. Independent samples t-test

Independent samples refer to two groups of data obtained by measuring different samples [8]. Independent samples test is used to deduce whether there is significant difference between two groups of data [9]. The experiment class and the control class were considered different samples. Therefore, in the pretest and posttest, the test scores of the experiment class and the control class were evaluated using independent samples t-test, and the mean value and difference were compared to infer whether the effect of oral English teaching and learning in the experiment class under the instruction mode supported by digital technology was significantly different from that in the control class under the traditional teaching mode without the support of digital technology [10].

4. Qualitative study of experimental data in oral English performance

4.1. Analysis of the total oral scores in the experimental class

As for the analysis of the total scores in the oral pretest, the statistical results are shown in **Table 1**.

Table 1. Statistics for oral pretest in the experimental class

N	Minimum	Maximum	Mean	Std. deviation	Std. error mean
50	9.0	12.0	10.11	.44136	.06242

Since the full score that can be obtained in the oral test was 15, and the mean score was 10.11, the students' oral English performance was in middle level.

However, according to **Table 2**, we can witness a remarkable rise in the students' oral English performance. The average score in the posttest was 12.29, which was 2.18 points higher than that of the pretest. Also, both the minimum score (11.0) and the maximum score were higher than the corresponding ones in the pretest.

Table 2. Statistics for oral posttest in the experimental class

N	Minimum	Maximum	Mean	Std. deviation	Std. error mean
50	11.0	14.0	12.29	.55227	.07810

Concerning the distribution of scores, the majority of students achieved scores ranging from 9.5 to 10.5 in the pretest, while most of them achieved scores from 12.0 to 12.5 in the posttest. From the distribution of their scores, significant growth was seen in the posttest.

The paired sample t-test in **Table 3** confirmed that there was a significant difference between the scores in the pretest and the posttest ($P = 0.000 < 0.05$). As a result of oral English blended teaching, the students' oral proficiency improved noticeably.

Table 3. Paired sample t-test for the oral pretest and posttest in the experiment class (full score = 15)

	Paired differences					t	d.f.	Sig. (2-tailed)
	Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
				Lower	Upper			
Pair 1 Posttest-pretest	1.2600	.74396	.10521	1.0485	1.4714	11.98	49	.000

4.2. Analysis of the oral scores in each module in the experiment class

In a separate evaluation of each model in the oral test, we discovered that the posttest scores of the three modules (accuracy and range of language, discourse coherence and length, as well as appropriateness and flexibility) were higher than the pretest scores.

Tables 4 and **5** showed a significant improvement in the performance of students in language accuracy and range as well as flexibility and appropriateness. Although there were still a few grammatical and lexical slips in the posttest, these slips had trivial effect on students' communication. Through their performance in spoken English, we could see that the students had a relatively wide store of vocabularies, and their pronunciation was comparatively good. Besides, in most parts, the students were willing to participate in discussions and express their ideas in appropriate language. In the pretest, however, there were quite a few grammatical mistakes in the usage of English. At times, these mistakes would hinder normal communication. Most students did not have sufficient vocabulary and grammatical structures and were unable to complete the given task or actively engage in discussions.

Table 4. Paired sample t-test of the pretest and posttest scores in grammatical range and accuracy (total score = 5)

Test	N	Mean	Std. Deviation	t	Sig. (2-tailed)
Posttest	50	4.36	.9115	12.05	.010
Pretest	50	3.90	1.4233		

Table 5. Paired sample t-test of the pretest and posttest scores in flexibility and appropriateness (total score = 5)

Test	N	Mean	Std. Deviation	t	Sig. (2-tailed)
Posttest	50	4.47	.7174	7.893	.002
Pretest	50	3.96	2.0843		

According to **Table 6**, there were significant differences in the pretest and posttest scores in terms of discourse length and coherence. As shown in the posttest results, more students had developed the ability to think and talk in the second language. They were able to quickly extract appropriate vocabulary and terms from their memory, thus showing desirable fluency in oral expression. Pauses were also significantly reduced, and they were observed to be more active in conversations. A number of students mentioned that they tend to apply Chinese thinking in English communication. This was the major reason they found it difficult to communicate in English. However, after a semester of lectures and practices, especially through many intensive language exchanges with their lecturers and other students under blended teaching, they began to break away from the Chinese thinking and express themselves better in English.

Table 6. Paired sample t-test of the pretest and posttest scores in discourse length and coherence (total score = 5)

Test	N	Mean	Std. Deviation	t	Sig. (2-tailed)
Posttest	50	4.27	.5284	5.371	.002
Pretest	50	3.98	.7718		

4.3. Analysis of the oral scores in the control class

As for the analysis of the scores in the oral pretest, the statistical results are shown in **Table 7**.

Table 7. Statistics for the oral pretest in the control class

N	Minimum	Maximum	Mean	Std. Deviation	Std. Error Mean
50	8.79	11.51	9.97	.42787	.06197

Since the full score that can be obtained in the oral test was 15, and the mean score was 9.97, the students' oral English performance was in middle level.

After one semester of teaching and learning oral English, the average score in oral English improved by 0.46 points from 9.97 to 10.43 (**Table 8**). Although the mean score increased, it was not significant.

Table 8. Statistics for the oral posttest in the control class

N	Minimum	Maximum	Mean	Std. Deviation	Std. Error Mean
50	8.91	12.02	10.43	.52144	.07213

It can be clearly seen in **Table 9** that there was no significant difference between the pretest and posttest results in the control class. The data showed that under conventional instruction, there was no significant improvement in the students' oral English skills.

Table 9. Paired sample t-test for the oral pretest and posttest in the control class (full score = 15)

	Paired differences				t	d.f.	Sig. (2-tailed)	
	Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
				Lower				Upper
Pair 2 Posttest-pretest	3.288	.3890	.945	1.0002	1.3433	1.112	49	.299

4.4. Comparison of the posttest oral scores between the experiment class and the control class

As shown in **Table 10**, the difference in the mean posttest oral scores between the experiment class and the control class was 1.87 points. By means of the test, a noteworthy difference was observed in the oral English proficiency between the two groups. As a result of the blended instruction mode, the oral English proficiency of the students in the experiment class improved noticeably, whereas the proficiency of the students in the control class showed no remarkable improvement. Therefore, the hypothesis that blended teaching supported by digital technology had evident advantages over traditional teaching in the development of university students' oral English performance is verified through this empirical study.

Table 10. Paired sample t-test for the posttest oral scores between the experiment class and the control class (full score = 15)

Test	N	Mean	Std. deviation	t	Sig. (2-tailed)
Experiment class	50	12.29	.5009	6.149	.001
Control class	50	10.42	.6875		

5. Conclusion

Through the integration of learning in- and out-of-class, oral English education was more systematic. Students generally held a positive attitude toward extracurricular oral English learning supported by digital technology. Although it took up their spare time, the majority of students thought it worthwhile. They believed that extracurricular learning, as a useful complement to in-class learning, would produce a synergistic effect.

Disclosure statement

The authors declare no conflict of interest.

Author contributions

Z.W. and N. B. conceived the idea of the study. Z.W. and Y.H. wrote the first draft of the paper. Z.W. and Q.T. revised the format of the article.

References

- [1] Wang Q, Castro D, 2010, Classroom Interaction and Language Output. *English Language Teaching*, 3(2): 175–186.
- [2] Meredith R, Catherine S, 2020, Analyzing Input Quality Along Three Dimensions: Interactive, Linguistic, and Conceptual. *Journal of Child Language*, 47(1): 5–21.
- [3] Allaw E, McDonough K, 2019, The Effect of Task Sequencing on Second Language Written Lexical Complexity, Accuracy, and Fluency. *System*, 85: 102–104.
- [4] Von S, Motha S, Kubota R, 2020, Race and Language Teaching. *Language Teaching*, 53(4): 391–421.
- [5] Rajesh V, Venkatesh L, 2019, Preliminary Evaluation of a Low-Intensity Parent Training Program on Speech-Language Stimulation for Children with Language Delay. *International Journal of Pediatric Otorhinolaryngology*, 122: 99–104.
- [6] Şahin M, Aybek E, 2019, Jamovi: An Easy to Use Statistical Software for the Social Scientists. *International Journal of Assessment Tools in Education*, 6(4): 670–692.
- [7] Mee W, Chua C, 1991, Regression Toward the Mean and the Paired Sample T Test. *The American Statistician*, 45(1): 39–42.
- [8] Gerald B, 2018, A Brief Review of Independent, Dependent and One Sample T-Test. *International Journal of Applied Mathematics and Theoretical Physics*, 4(2): 50–54.
- [9] Rietveld T, van Hout R, 2015, The T Test and Beyond: Recommendations for Testing the Central Tendencies of Two Independent Samples in Research on Speech, Language and Hearing Pathology. *Journal of Communication Disorders*, 58: 158–168.
- [10] Krzywinski M, Altman, N, 2014, Comparing Samples – Part I: Robustly Comparing Pairs of Independent or Related Samples Requires Different Approaches to the T-Test. *Nature Methods*, 11(3):

215–217.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.