International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 22s, 2025
https://www.theaspd.com/ijes.php

Investigating The Impact Of AI-Driven Personalized Learning Platforms On Vocabulary Acquisition Among Chinese EFL University Students

[1]Qingqing Cui, [2]Supyan Hussin
National University of Malaysia, cuiqingqingbtc@gmail.com
[2]National University of Malaysia, 1227420625@qq.com

INTRODUCTION

1. BACKGROUND OF THE STUDY

English is widely recognized as the global lingua franca, playing a critical role in international communication, trade, and education. For non-native speakers, acquiring proficiency in English is essential for accessing global opportunities and achieving success in academic, professional, and social contexts (Rao, 2019). Vocabulary acquisition is one of the core components of English language learning, directly influencing learners' integrated skills in listening, speaking, reading, and writing. Vocabulary is the foundation of language communication. To comprehend reading materials, non-native speakers must master 98% of the vocabulary within a text (Laufer & Ravenhorst-Kalovski, 2010).

As the primary vehicle of meaning in language, vocabulary serves as the medium for cognitive activities and facilitates communication through the construction of sentences. Proficiency in a language requires a robust and extensive vocabulary base, and the ability to memorize and comprehend vocabulary is a key indicator of effective English language learning (Chen, 2012). However, traditional vocabulary learning methods, such as rote memorization and word list drills, are often criticized for being monotonous and lacking real-world application (Yu, 2018). In the context of rapid globalization, English has become increasingly important as a global lingua franca. In China, English learning is a significant component of the education system. However, vocabulary acquisition remains a common challenge for Chinese university students. Insufficient vocabulary not only hinders students' reading, writing, listening, and speaking abilities but also limits their academic and professional development.These methods frequently result in low retention rates and limit learners' ability to apply acquired vocabulary in academic and communicative contexts. Enhancing the efficiency and effectiveness of vocabulary acquisition remains a critical issue in educational research. In recent years, the rapid advancement of Artificial Intelligence (AI) has introduced new possibilities for vocabulary learning through AI-driven personalized learning platforms, advancements in artificial intelligence (AI) have created new opportunities for language learning. AI-driven personalized learning platforms, through data analysis and machine learning algorithms, can customize learning paths based on learners' proficiency levels, interests, and study habits. This approach goes beyond the limitations of traditional classroom teaching and more effectively meets the individual needs of learners.Such platforms employ machine learning algorithms and data analytics for smart and adaptive learning systems, which adjust learning materials according to learners' behaviors and performance, creating personalized learning for each user (Rukiati, 2023). Meanwhile, gamification-young people learn better-by playing and motivating them through tasks and reward mechanisms (Dehghanzadeh, 2021). Spaced repetition techniques, on the other hand, based on the forgetting curve described by Ebbinghaus, have also been demonstrated to be effective in enhancing long-term vocabulary retention (Herzallah, 2019). Therefore, the addition of these features makes AI-driven personalized learning platforms a technology that can boost vocabulary acquisition process.

While existing studies have shown positive effects of AI-driven tools in areas such as grammar instruction and pronunciation training, research on their impact on vocabulary acquisition is relatively scarce. Specifically, how AI technology can enhance vocabulary learning among

Chinese university students has not been systematically explored. Therefore, this study focuses on investigating the impact of AI-driven personalized learning platforms on vocabulary acquisition among Chinese EFL (English as a Foreign Language) university students, as well as its potential educational implications.Despite the fact that there exists an increasing interest in AI tools in higher education, research specifically on the impact of AI-enabled platforms on vocabulary learning among Chinese EFL students is still new. Not only that, but the relative effectiveness of the different functional units of the platforms has not yet been undeniably confirmed with proper empirical studies. For this reason, the current study aims to fill this research gap by rigorously assessing AI-driven personalized learning platforms in vocabulary acquisition and by exploring the key features that account for success in learning outcomes.

2 PROBLEM STATEMENT

Vocabulary building is a burning issue in the curriculum of English as a Foreign Language for Chinese university students. Nevertheless, many of them still find it challenging. Traditional learning methods more often than not prove to be ineffective in engaging students and upholding personalization. Consequently, learners are usually burdened with low retention standards, and regrettably, the already acquired vocabulary is hardly applicable when it comes to advanced academic studies or real-life corporate scenarios (Xu, 2014). Though AI-driven platforms for personal learning in education are now experiencing massive popularity, their real efficacy as regards vocabulary acquisition among EFL learners in the relatively controlled environment of Chinese higher education is largely unknown. Furthermore, happenings like Quizlet, Memrise, and DenoPark take up numerous features such as adaptive learning, gamification, and spaced repetition, for instance; however, there still is no evidence as to which of these features is the strongest factor in the overall learning outcomes (Nguyen, 2022). The rise of AI-driven personalized learning platforms offers a potential solution to these challenges. By utilizing advanced algorithms and data analytics, these platforms can adapt to individual learners , abilities, providing tailored content, immediate feedback, and engaging learning experiences. However, while studies have shown the general benefits of AI in education, there is limited empirical research specifically examining its impact on vocabulary acquisition among EFL learners, particularly in the Chinese university context.Employee engagement takes up a significant position as it shows the importance of examining the differential effects of the subcomponents for optimal design and implementation.

For educators and educational institutions, the lack of research evidence is one of the key difficulties while often trying to incorporate these instruments into traditional language teaching. In view of this, the current research aims to cover two major aspects, which are: (1) assessing the overall effectiveness of AI-powered platforms in learning words and (2) establishing the causes for which the result is that each learning outcome is determined by a marked degree. By fulfilling this purpose, the current paper will serve teachers, lenders, and decision-makers.

3 RESEARCH OBJECTIVES

To evaluate the overall effectiveness of AI-driven personalized learning platforms in improving vocabulary acquisition among Chinese EFL university students.

To examine the impact of specific platform features (adaptive learning, gamification, and spaced repetition) on vocabulary learning outcomes.

To investigate whether the use of the platform enhances learners' motivation and engagement.

To analyze the differences in vocabulary acquisition performance among students with varying English proficiency levels.

4 Research Questions

Does the AI-driven personalized learning platform significantly improve vocabulary acquisition among Chinese EFL university students?

How do different platform features (adaptive learning, gamification, and spaced repetition) affect vocabulary learning outcomes?

Does using the platform significantly enhance learners' motivation and engagement?
What are the differences in vocabulary acquisition performance between students with different English proficiency levels?

5 RESEARCH METHODOLOGY

5.1 RESEARCH DESIGN

This study adopts a quasi-experimental design with a pre-test and post-test structure to evaluate the effectiveness of AI-driven personalized learning platforms in improving vocabulary acquisition among Chinese EFL university students. The design includes multiple experimental groups using distinct features of the AI platform and a control group employing traditional learning methods. Participants are further stratified by their English proficiency levels to investigate performance differences.

The study incorporates three key aspects:

Evaluation of platform effectiveness: Comparison between experimental groups and the control group to assess the overall impact of the AI-driven platform on vocabulary acquisition.

Feature-specific analysis: Examination of the relative effectiveness of three distinct platform features: adaptive learning, gamification, and spaced repetition.

Proficiency-level analysis: Exploration of differences in vocabulary acquisition performance between low-proficiency and high-proficiency learners.

More so the current research combines a survey design with pre-test and post-test analysis for monitoring learning motivation and engagement changes over time. Therefore, the primary goal of this study is to involve standardized survey instruments to widely collect information on the motivation and engagement levels of the students in the experimental and control groups. The strategy here is to expose how the AI-driven platform leads to an increase in the number of times the students engage in the learning process during the intervention.

5.2 THE POPULATION AND SAMPLING

The term "study population" is used to signify a specific, defined group of individuals or entities that possess a given characteristic or are associated with a given attribute (Buchstaller & Khattab, 2013).

In this study, the population under scrutiny consisted of undergraduate students from Jinan University, which is located in Shandong Province, China. Jinan University occupies a key status in the local community as an important university with a multicultural and interdisciplinary student body. According to recently published data by the administration of the university, the number of undergraduate students currently exceeds 30,000. This group is quite diverse in terms of the academic and personal backgrounds, thus being the best fit to conduct an experiment on the influence of AI-based individualized learning solutions on vocabulary development. Besides the purpose of this study, the participants were drawn from this population to guarantee the achievement of the research goal.

The participants were first-year undergraduates from Shandong Normal University in Shandong Province, China. To ensure that selected participants met the research goals, the following criteria were formulated. For starters, every participant was an undergraduate undertaking EFL (English as a Foreign Language). Moreover, participants were required to display a basic level of English knowledge to be able to actively engage with assigned vocabulary exercises implemented by the adaptive learning technology. In addition to these, participants needed to be enrolled for the12-week therapy, to take both pre-and post-tests, and to fill in an online survey regarding their behavioral tendencies. The participation of the experiment was on a voluntary basis, with all subjects signing informed consent forms. Given that they were participating in the study voluntarily, students were kept informed regarding their rights. In addition, they had the right to leave the study any time with no adverse impact on their academic standing or their other rights. Thereby, the selection criteria were meant to endorse the relevance of the participants to the research and this compliance with the ethics of the research process.

Reference will be made here to students with similar AI platform experience or to those having intermediate knowledge of the platforms' uses, in order to mitigate potential bias linked to

previous familiarity. Students who did not comply with the pre-test instructions for ushering, such as filling out the required documents, were not part of the experiment formalization process. Likewise, those who were either unable to continue with the experimental program or had to withdraw from it due to conditions beyond their control, where necessary, would not be included in the data assessment.

The stratified random sampling method was used in the course of this research project to increase the definitiveness of the outcomes drawn and to attain scientific accuracy. The whole students ' demographic of Shandong Normal University, estimated about 30,000 according to the university records, was considered to be the study population. To find out if English proficiency level affected the outcomes of learning, national English exams were used to stratify the students into two proficiency levels: a low proficiency level group (CET-4 and below) and a high proficiency level group (CET-6 and above). Among students with the same proficiency level, sampling randomly was used to choose 120 students and thus there were 240 participants of equal gender distribution. Those participants were subsequently split into one of four groups, one using adaptive learning techniques, one using gamification techniques, one using spaced repetition techniques, and lastly, the control group in a remainder of participants. The sample distribution provided equal representation of each proficiency level and controlled for bias during the randomization process.

Grouping Strategy		
proficiency levels	Low-Proficiency Group	Students with CET-4 level or below
	High-Proficiency Group	Students with CET-6 level or above
Within each proficiency level, participants will be further randomly assigned into four subgroups(30 students per proficiency level)		
Experimental Group 1:	Using the adaptive learning feature of the AI platform(low:30;High:30)	
Experimental Group 2	Using the gamification feature of the AI platform(low:30;High:30)	
Experimental Group 3	Using the spaced repetition feature of the AI platform (low:30;High:30)	
Control Group	Using traditional vocabulary learning methods, such as word lists and written exercises(low:30;High:30)	

5.3 DATA COLLECTION METHOD

5.3.1 SURVEY

Instrument:

College Students' Learning Motivation Questionnaire

The College Students' Learning Motivation Questionnaire was made by Tian Lan and Pan Weigang in 2006 and is based on Ausubel's three-factor motivational theory. According to Ausubel, achievement motivation in a school context should

include at least three driving forces: cognitive drive, self-enhancement drive, and affiliation drive. In the development of this questionnaire, the learning motivation of college students was defined as the "virtual or realistic reasons that stimulate college students' learning." The key motivational components assessed include satisfying the need for knowledge, developing professional interests, enhancing abilities, gaining recognition and praise from others, creating employment opportunities, and serving society.

This self-reported questionnaire consists of four dimensions: interest in knowledge acquisition, pursuit of ability, reputation acquisition, and altruistic orientation, with a total of 34 items. It adopts a five-point Likert scoring method: "strongly agree" is scored 5, "somewhat agree" is scored 4, "somewhat disagree" is scored 3, "mostly disagree" is scored 2, and "strongly disagree" is scored 1.

The questionnaire demonstrates excellent reliability and validity. Regarding reliability, the overall internal consistency coefficient (Cronbach's α) is 0.9003, while the coefficients for each subscale range from 0.7719 to 0.8389, indicating high internal consistency for the questionnaire and its subscales. Additionally, all item discrimination indices exceed 0.35, demonstrating strong discriminatory power for the items. Regarding validity, principal component analysis extracted four factors (interest in knowledge acquisition, pursuit of ability, reputation acquisition, and altruistic orientation), which together explained 45.152% of the total variance. Factor loadings ranged from 0.425 to 0.782, further supporting the structural validity of the questionnaire. Furthermore, the correlations between subscales and the overall scale ranged from 0.741 to 0.888, and correlations among subscales ranged from 0.233 to 0.548, indicating that the subscales effectively reflect the measurement objectives of the questionnaire while maintaining a certain degree of independence. The questionnaire exhibits excellent reliability and validity, making it a robust tool for assessing college students' learning motivation(Tian & Pan, 2006).

Utrecht Work Engagement Scale-Student (UWES-S)

Schaufeli and colleagues (2002) developed the Utrecht Work Engagement Scale-Student (UWES-S) using university students as the sample population. For university and graduate students, studying is a highly professional activity, particularly in research work, which shares many characteristics with workplace tasks. Study Engagement refers to a persistent, positive, and fulfilling emotional and cognitive psychological state related to study, research, and employment. It is comprised of three core dimensions: Vigor, Dedication, and Absorption. Fang Laitan and colleagues (2008) revised the UWES-S scale, which consists of 17 items (e.g., "I feel energized while studying"), covering the three dimensions of vigor, dedication, and absorption. The scale uses a seven-point Likert scoring method, with higher total scores indicating greater levels of study engagement. In Jiang Yuan et al.'s (2020) study, The scale exhibited remarkable reliability; 0.96 is the general Cronbach's α coefficient, and the dimension-specific Cronbach's α coefficients range from 0.89 to 0.93.

Administration:

This study's survey methodology consists of two critical sections: the initial questionnaire to establish a baseline before the intervention and the final post-test questionnaire at the end of the program, allowing for the detection of changes in learning motivation and engagement over time.

The pre-test questionnaire is performed in the first week to determine the participants' levels of bringing up their motivation and engagement to learn. This will be a reference point for the remaining courses. Introducing the researchers to their instruments is the next step in the distribution process. The participants are provided with instructions on how to correctly complete the questionnaire and are also made aware of the different factors and matters to consider regarding responding to the questions. The questionnaire is filled in either by participants printing and filling it in or by submitting it online through the Questionnaire Star platform; this is to ensure that all participants can answer it comfortably and with good data.

The post-test questionnaire will be conducted in the 12th week on the last day of the intervention, determining the impact of the intervention on all the researched variables. The process mirrors that of the pre-test to ensure data comparability. The research team notifies participants of the

questionnaire arrangements in advance and provides full guidance during completion to minimize misunderstandings. Whether completed online or offline, the questionnaires must be finished within the specified timeframe. Researchers then review the completeness and validity of the responses, removing any clearly invalid entries. Once collected, offline questionnaires are entered into electronic spreadsheets, and online responses are exported as structured data. All data are then systematically organized, cleaned, and archived.

5.3.2 INTERVENTION DESIGN

Duration

The intervention for this experiment lasts for 12 weeks, providing sufficient time to observe changes in vocabulary acquisition.

Learning Schedule

Experimental Group

Each participant in the experimental group uses a designated AI-driven learning module for 30 minutes daily over the course of 12 weeks. The learning modules focus on three core features: adaptive learning, gamification, and spaced repetition. Each participant is assigned to use only one specific module to evaluate its individual effectiveness. The platforms employed to carry out this endeavor are Quizlet, Memrise, and DenoPark, all of which utilize AI-driven personalized learning features:

Quizlet: It has an adaptive pathway. Quizlet uses machine learning to detect students' learning flow and tailor content based on their proficiency, error patterns, and progress through the content. The system is, thus, able to intensify the frequency of practice of the words that are commonly misstated and diversify the mode of exercise (i.e., spelling, multiple-choice, and fill-in-the-blank) to foster retention among users. This is, in part, because it suggests sets of words or tasks to show those words that need attention and to guide students toward a more efficient process.

Memrise: Features a feature inspired by spaced repetition. Memrise optimizes the consideration of the Ebbinghaus forgetting curve by reviewing work intervals that have been spaced out. With new vocabularies and words of the day, the system is also recommended to be reviewed (e.g., after one day, three days, one week). As it is done, words proceed from being in short-term memory to permanent memory. Also, the program has an adaptive system that adds/updates the review frequencies based on the performance of students (in the form of correct or incorrect answers). Mastered common words are reviewed, redundantly enhancing time efficiency.

DenoPark: The unique feature of this platform is that it adds an element of gamification to language learning to increase user involvement and increase the effectiveness of training. Its core features are an RPG game mode where users can combat with monsters and raise upgrades after completing spelling tasks. This role-playing is an extension of the overall fun experience. The platform has a health points (HP) scheme as well. It will be like a countdown clock where wrong answers will lead to HP reduction; hence, the challenge will be increased, making learning fun and motivational. **Control Group**

Participants in the control group study for the same amount of time (30 minutes daily) using traditional methods, such as memorizing word lists and completing paper-based exercises. The learning time and frequency are kept consistent with the experimental group.

Administration:

This study collects data through two methods to comprehensively evaluate learning outcomes and changes in learning behavior. Standardized vocabulary tests (TOEFL iBT vocabulary question bank: covering commonly used vocabulary in academic reading and listening contexts) are used in Week 1 (pre-test) and Week 12 (post-test) to assess participants' vocabulary proficiency. The test content includes:

High-frequency vocabulary (30%): Tests common words such as "basic" and "important" to evaluate participants' foundational vocabulary skills.

Academic vocabulary (50%): Includes words used in academic contexts, such as "hypothesis" and "analyze," to assess academic application abilities.

Low-frequency vocabulary (20%): Tests rare but meaningful words, such as "obscure" and "ephemeral," to evaluate deep memory and vocabulary expansion.

The test consists of multiple-choice questions, fill-in-the-blanks, matching, and short-answer questions to assess vocabulary retention and practical application. The pre-test and post-test use different questions but maintain consistent difficulty and structure to ensure scientific validity and comparability. Test results are scored based on accuracy to quantify learning outcomes. Additionally, the study records participants' learning behaviors through platform logs, including daily study time, task completion rate, error rate, and review frequency. For instance, Quizlet tracks study time and completed flashcard sets, Memrise records word review frequency and accuracy, and DenoPark logs RPG game performance and review effectiveness. These log data are automatically generated and exported by the platforms for analyzing the relationship between learning behaviors and vocabulary acquisition.

6 DATA ANALYSIS

This study utilized SPSS 27.0 software to analyze the data collected through survey and test. The responses were first cleaned and organized to ensure data quality, with variables numerically coded for statistical analysis. Descriptive statistics were calculated to summarize participants' demographic information and overall response trends. Reliability analysis, using Cronbach's α, assessed the internal consistency of the questionnaire items, while exploratory factor analysis verified the construct validity. Inferential statistical methods, including t-tests, ANOVA, and correlation analysis, were employed to examine group differences and relationships among variables, such as motivation, engagement, and vocabulary acquisition.

Research Questions	Statistics Test
Does the AI-driven personalized learning platform significantly improve vocabulary acquisition among Chinese EFL university students?	Paired-samples t-test: Compare pre-test and post-test vocabulary scores within the same group to determine improvements.
How do different platform features (adaptive learning, gamification, and spaced repetition) affect vocabulary learning outcomes?	One-way ANOVA: Compare post-test vocabulary scores among three experimental groups using different platform features.
Does using the platform significantly enhance learners' motivation and engagement?	Paired-samples t-test: Compare pre-test and post-test motivation and engagement questionnaire scores for significant changes.
What are the differences in vocabulary acquisition performance between students with different English proficiency levels (low and high)?	Independent-samples t-test: Compare post-test vocabulary scores between low and high proficiency groups.

7 VALIDITY AND RELIABILITY

To ensure the validity of this study, multiple strategies are implemented to verify the accuracy and relevance of the tools and methods. The questionnaire and standardized vocabulary tests are based on established frameworks, such as the TOEFL and IELTS vocabulary sections, ensuring their alignment with the constructs being measured. Expert reviews improve the precision of the questionnaire items in order to achieve content validity. This is followed by exploratory factor analysis (EFA), a technique that validates construct validity by checking item alignment on theoretical dimensions such as motivation and engagement. Likewise, criterion-related validity is assessed with the analysis of correlations of post-test vocabulary results, platform log data (like time and task completion), and self-reported questionnaire results. As a result, consistency across measures is ensured.

Reliability is ensured by evaluating the internal consistency of the questionnaire using Cronbach's α, where a value above 0.7 indicates an acceptable reliability standard. The test-retest reliability of the participants is performed on a small group of people to measure the stability of the questionnaire over the time period. Different standardized vocabulary tests (such as TOEFL and IELTS) are applied to guarantee similar and accurate indices of vocabulary development. Along with this, platform log data is automatically recorded, and the AI platforms export it to protect accuracy and help decrease human error. The data cleaning and cross-verification processes give, in addition, the reliability and integrity of all the collected data, resulting in only the authentic data used for research, and hence good and quality findings.

8 REFERENCES

1. *Zhonggen, Y. (2018). Differences in serious game-aided and traditional English vocabulary acquisition. Computers & Education, 127, 214-232.*
2. *Rukiati, E., Wicaksono, J. A., Taufan, G. T., & Suharsono, D. D. (2023). AI on learning English: Application, benefit, and threat. Journal of Language, Communication, and Tourism, 1(2), 32-40.*
3. *Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., & Noroozi, O. (2021). Using gamification to support learning English as a second language: a systematic review. Computer Assisted Language Learning, 34(7), 934-957.*
4. *Rao, P. S. (2019). The importance of English in the modern era. Asian Journal of Multidimensional Research (AJMR), 8(1), 7-19.*
5. *Herzallah, R.(2019). Investigating the Efficiency of Using Spaced Repetition Technique in Enhancing Long-term Propositional Knowledge Retention: The Case of Master One Students at the Section of English at Mohamed Kheider University of Biskra (Doctoral dissertation).*
6. *Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners , vocabulary size, and reading comprehension. Reading in a Foreign Language, 22(1), 15-30.*
7. *Chen, X. H. (2012). Technology-supported junior high school English vocabulary teaching. Journal of Inner Mongolia Normal University (Educational Science Edition), 25(6), 125-127.*
8. *Nguyen, T. T., Nguyen, D. T., Nguyen, D. L. Q. K., Mai, H. H., & Le, T. T. X. (2022). Quizlet as a tool for enhancing autonomous learning of English vocabulary. AsiaCALL Online Journal, 13(1), 150-165.*
9. *Xu, G. (2014). A study on English vocabulary learning strategies used by Chinese college students. Studies in Literature and Language, 9(1), 111.*
10. *Buchstaller, I., & Khattab, G. (2013). Population samples. Research methods in linguistics, 74-95.*
11. *Tian, L., & Pan, W. G. (2006). Development of the College Students , Learning*