

PiBAL-ML: A Machine Learning-Based Personalized Listening Platform for EFL Learners

Kübra OKUMUŞ DAĞDELER^{1*}  Yasin GÖRMEZ¹  Samet Çağrı KIZKAPAN¹ 

¹ Sivas Cumhuriyet University, Sivas, Türkiye.

ABSTRACT

Individuals exhibit variations in their needs, interests, and affective behaviours during the language learning process. Recognizing and accommodating these differences is crucial for optimizing language acquisition. Advancements in technology have provided numerous tools and opportunities to tailor language learning experiences to individual preferences, thereby facilitating a more effective learning journey. Regarding these rationales, this study aimed to propose a personalized language learning platform which was called Personal Interest Based Listening- Machine Learning (PiBAL-ML) and then determine its effect on English as Foreign Language (EFL) learners' listening skills. Thus, the study unfolded in two distinct phases: the design phase and the testing phase. The PiBAL-ML system, at its core, featured a diverse array of listening materials carefully curated to align with the learners' specific fields of interest. Furthermore, it included a variety of question types, and the number of questions was dynamically adjusted based on participants' performance and progress. The effectiveness of the system was examined through an explanatory sequential research design in which a quasi-experimental treatment was implemented over one semester. Data were gathered using a listening achievement test and an open-ended questionnaire. The participants consisted of 44 EFL students enrolled at a state university. The quantitative results revealed that there was no significant difference in overall scores between the control group and the experimental group. However, experimental group outperformed control group in multiple-choice test questions. Lastly, the study highlighted that the learners were satisfied with the system as it presented different activities and texts related to their interests. Based on these findings, the study concluded that machine learning based personalization could enhance foreign language learning.

Keywords: Adaptive technologies, English as a Foreign Language, listening skill, machine learning, personalized learning.

* Corresponding author's email: kokumus@cumhuriyet.edu.tr

Cite as: Dağdeler-Okumuş, K., Görmez, Y., & Kızıkan, S.Ç. (2026). PiBAL-ML: A Machine Learning-Based Personalized Listening Platform for EFL Learners. *International Journal of English for Specific Purposes*, 4(1), 44-62. <https://doi.org/10.70870/joinesp.1922933>

PiBAL-ML: Yabancı Dil Olarak İngilizce Öğrenenler İçin Makine Öğrenmesi Tabanlı Kişiselleştirilmiş Dinleme Platformu

ÖZ

Bireyler, dil öğrenme süreci boyunca ihtiyaçları, ilgi alanları ve duygusal davranışları açısından farklılıklar sergilerler. Bu farklılıkları fark etmek ve bunlara uyum sağlamak, dil edinimini daha verimli hale getirmek açısından hayati önem taşır. Teknolojideki gelişmeler, dil öğrenme deneyimlerini bireysel tercihlere göre uyarlama imkânı sunan çok sayıda araç ve fırsat sağlamış, böylece daha etkili bir öğrenme süreci oluşturulmuştur. Bu gerekçeler doğrultusunda, bu çalışma, Kişisel İlgi Temelli Dinleme-Makine Öğrenmesi (PiBAL-ML) adı verilen kişiselleştirilmiş bir dil öğrenme platformu önermeyi ve ardından bunun Yabancı Dil olarak İngilizce öğrenenlerin dinleme becerileri üzerindeki etkisini belirlemeyi amaçlamıştır. Bu nedenle, çalışma iki ayrı aşamada gerçekleştirilmiştir: tasarım aşaması ve test aşaması. PiBAL-ML sistemi, öğrenenlerin belirli ilgi alanlarına uyacak şekilde özenle seçilmiş çok çeşitli dinleme materyallerini içermektedir. Ayrıca, içerdiği çeşitli soru türleri ve soru sayısı, katılımcıların performansına ve ilerlemesine göre dinamik olarak ayarlanmıştır. Sistemin etkililiği, bir sömestr boyunca yarı deneysel bir uygulamanın gerçekleştirildiği açıklayıcı sıralı araştırma tasarımıyla incelenmiştir. Veriler, dinleme başarı testi ve açık uçlu bir anket kullanılarak toplanmıştır. Çalışmanın katılımcıları, bir devlet üniversitesinde kayıtlı 44 Yabancı Dil olarak İngilizce öğrencilerdir. Nicel sonuçlar, kontrol grubu ile deney grubu arasında genel puanlarda anlamlı bir fark olmadığını ortaya koymuştur. Ancak, deney grubu çoktan seçmeli test sorularında kontrol grubundan daha iyi performans göstermektedir. Son olarak, çalışma, öğrencilerin ilgi alanlarıyla ilgili farklı etkinlikler ve metinler sunan sistemden memnun olduklarını vurgulamaktadır. Bu bulgulara dayanarak, çalışma, makine öğrenmesi tabanlı kişiselleştirmenin yabancı dil öğrenimini geliştirebileceği sonucuna varılmıştır.

Anahtar Kelimeler: uyarlanabilir teknoloji, yabancı dil olarak İngilizce, dinleme becerisi, makine öğrenmesi, kişiselleştirilmiş öğrenme

1. Introduction

Over years, technological devices such as radios, tape recorders, and CDs have played a significant role in enhancing listening skills. As technology has advanced, particularly in the realm of educational technology, listening websites and mobile applications have emerged, enabling learners to engage independently without direct teacher guidance. Many of these tools incorporate adaptive systems, providing a degree of personalized learning. However, teaching listening remains a persistent challenge in language education despite these advancements. One major issue is the passive nature of traditional listening activities, where learners often listen without actively engaging in comprehension processes. Additionally, listening comprehension is influenced by various factors, including speech rate, accents, background noise, and the complexity of spoken discourse, making it difficult for learners to develop strong listening skills using generic, one-size-fits-all approaches. Moreover, while existing digital tools offer some level of interactivity, they often lack the ability to adapt dynamically to individual learners' needs, particularly in real-time listening comprehension development. Many students struggle with maintaining motivation and engagement when using such tools, as they do not provide immediate, personalized feedback or scaffolding to address learners' specific listening difficulties. Furthermore, teachers frequently face constraints in providing individualized instruction, given the limitations of classroom time and the diverse proficiency levels of students.

Given these challenges, it is crucial to explore more sophisticated solutions that can enhance listening skill development. Personalized learning can play an important role in these solutions by assisting language acquisition and skill mastery. If personalized learning is not effectively integrated into instruction, students may struggle to achieve high proficiency levels (Looi et al., 2009). Based on these

rationales, this paper proposes a model that leverages machine learning to facilitate personalized listening skill development. The concepts of personalized learning (PL, hereafter) and machine learning are explored, followed by a review of studies related to personalized language learning (PLL, hereafter). Subsequent sections include methodology, findings, discussion, and conclusion, highlighting the potential of machine learning-driven approaches in addressing the challenges of listening skill instruction.

1.1. Personalized language learning

Although the concept of personalized learning (PL) is not widely understood (Basham et al., 2016) and lacks a single, universally accepted definition (Bulger, 2016), several attempts have been made to conceptualize it. Traxler (2009) defines PL as “*learning that recognizes diversity, difference, and individuality in the ways that learning is developed, delivered, and supported*” (p. 17). Similarly, the U.S. Department of Education, Office of Educational Technology (2017) characterizes personalized learning as “*instruction that is adapted to learners’ individual pace, needs, and interests, allowing flexibility in learning objectives, content, and sequencing*” (p. 1). From a learner-centred perspective, personalization enables students to engage with learning across different contexts, monitor their own progress, and access educational resources aligned with their skills and abilities (Petersen et al., 2009). In this respect, individual differences, personal needs, and learner development constitute the core components of PL (Peng et al., 2019). Given that learners differ in terms of learning styles, learning speed, and motivation levels, the language learning process must account for these variations.

Advancements in technology have substantially facilitated and accelerated the personalization of learning. For instance, LLM-based AI-enabled educational systems or environments powered by multiple agents are capable of analysing students’ progress in real time through continuous learner interactions, enabling both teachers and AI agents to dynamically adapt instructional content to address learners’ specific needs (Sharma et al., 2025). Today, numerous online applications and platforms support language education. However, despite the abundance of available resources, identifying appropriate materials remains challenging. Even when resources are accessible, they may not adequately address individual learner needs. Consequently, recommending suitable resources that support personalized learning from a vast pool of options has become a critical concern (Wang, 2023). Despite this need, research on adaptive technologies designed to support personalized English language learning remains limited (Kerr, 2016). The research gap identified by Kerr has not been sufficiently addressed in subsequent years, with relatively few follow-up studies conducted. Moreover, as technology evolves rapidly, personalization approaches also continue to change. Therefore, research on technology-mediated personalization must be continuously updated to remain relevant. Although the literature includes studies advocating personalized learning environments, this study, like some previous work, proposes a system and reviews existing systems to highlight differences in personalization approaches.

Personalised learning and adaptive learning are two related yet distinct concepts that are increasingly shaping modern education. Although adaptive learning and personalized learning are sometimes used interchangeably, personalized learning focuses on tailoring the pace and approach of instruction to individual student needs, while adaptive learning uses technology and real-time data to automatically adjust instruction as students progress (Adams Becker et al., 2017). The distinction between the two lies in their scope and emphasis. Personalized learning is considered the broader of the two concepts, as it goes beyond content delivery adjustments to encompass the whole learner. Walkington (2013) argues that personalized learning extends beyond simply adjusting content delivery, as it encompasses students’ individual interests, motivations, and real-world experiences to create more meaningful learning

opportunities. Similarly, Pane et al. (2015) describe personalized learning as addressing multiple dimensions of the learner, including academic strengths, interests, and personal goals. It is for this reason that this study preferred the term personalized learning, even though it incorporates adaptive learning technologies as part of its instructional approach.

Research on Personalized Language Learning (PLL) has predominantly focused on mobile-based systems (Chen & Chung, 2008; Chen & Hsu, 2008; Hsu et al., 2013; Petersen et al., 2009) and web-based platforms, with relatively few review studies available. A notable concentration of studies emphasizes Mobile-Assisted Language Learning (MALL), largely due to the affordances mobile technologies offer for personalization. Chen et al. (2021) identified several key research areas within PLL, including personalized grammar learning, PLL networks, recommendation systems, game-based PLL, personalized feedback, the role of anxiety in PLL, and mobile-assisted PLL, with the latter receiving the greatest research attention. This close relationship between PL and mobile learning is evident in multiple studies. For example, Hsu et al. (2013) developed a mobile recommendation-based learning system and reported significant learning gains in the experimental group. Likewise, Chen and Hsu (2008) designed an intelligent mobile learning system for reading skills, integrating Chen and Chung's (2008) Item Response Theory-based personalized vocabulary system, which led to improvements in both reading comprehension and vocabulary acquisition. Other studies further demonstrate the role of mobile technologies in personalization. Wu et al. (2014) employed RFID-based location-aware technology combined with a portfolio system, resulting in enhanced reading comprehension and learner engagement. Petersen et al. (2009) introduced PALLAS, a mobile-based personalized learning system that teachers evaluated positively for its contribution to lifelong learning. Nevertheless, although "anytime and anywhere" learning (Kukulska-Hulme & Shield, 2008) is a defining feature of MALL, personalized learning extends beyond mobile contexts to encompass a wider range of methods and platforms.

Beyond mobile technologies, researchers have also examined web-based PLL systems using approaches such as data mining, parse key trees, text algorithms, and multimodal conversational agents to generate personalized recommendations. Zou (2011) applied data mining techniques to support automated recommendations in web-based autonomous language learning and demonstrated their effectiveness over a semester-long implementation. Fang et al. (2017) proposed a parse key tree model that improved grammar question recommendations compared to traditional systems. Similarly, Griol et al. (2014) developed a multimodal conversational agent for personalized grammar, vocabulary, and listening practice, reporting increased user satisfaction and improved learning outcomes.

Several studies have specifically focused on personalized vocabulary learning. Kwon et al. (2010) designed a retrieval-based personalized vocabulary system that significantly enhanced learners' vocabulary knowledge. Jung and Graf (2008) proposed a web-based vocabulary learning approach based on personalized word games, incorporating lexicon assessments and association tasks, and evaluated learning outcomes through multiple-choice tests. More recently, Yang and Ogata (2023) integrated repeated retrieval practice into a personalized e-book environment, which resulted in higher reading achievement and behavioural engagement compared to non-personalized versions. In the area of listening, Jiang (2025) proposed and evaluated a deep-learning-based Convolutional Neural Network Path Recommendation (CNN-PR) algorithm for personalized English listening learning. The findings indicate that the algorithm effectively recommends listening materials tailored to learners' vocabulary complexity, proficiency levels, reading speed, and topic relevance. Overall, the CNN-PR system was successful in selecting relevant and engaging listening resources, leading to improved comprehension, increased learner motivation and satisfaction, and more efficient management of study time.

Although these studies employ diverse personalization techniques, they share a common objective of enhancing personalized learning experiences, with many reporting positive learning outcomes. However, the existing PLL literature largely concentrates on reading, vocabulary, and grammar, with comparatively limited attention to listening skills, apart from Jiang (2025) and the work of Griol et al. (2014) that addressed multiple language domains. To address this gap, the present study aims to enhance EFL learners' listening skills through machine-learning-based personalization, an approach that remains underexplored in personalized language learning research. In addition, Bernal Parraga et al. (2025) emphasize that artificial intelligence tools play a substantial role in shaping individualized learning in English language instruction.

1.2. Machine learning in language learning

Machine learning has been successfully applied across various fields, offering solutions to complex, non-linear problems. In personalized learning (PL), machine learning is frequently used, though research remains limited. Xu et al. (2020) developed a web-based learning support system using big data mining to track student progress and provide personalized assistance in distance learning. Zhang (2021) proposed a machine learning-based English learning platform in China, demonstrating that learners with similar backgrounds and goals may require different materials. A recommendation system was developed to address existing issues in English language education. Several studies have leveraged machine learning for language learning personalization. For instance, Pokrivcakova (2019) distinguished between adaptive and intelligent educational systems, emphasizing the need for more efficient teaching tools. Ross et al. (2013) used RGB-D sensors and k-means clustering to track student attention, classifying behaviours via support vector machines. Troussas et al. (2015, 2018, 2019) developed AI-driven recommendation systems integrating fuzzy logic for personalized English and French learning, detecting spelling, tense, and auxiliary verb errors. Yeung and Lee (2018) proposed a text retrieval algorithm to recommend reading materials based on vocabulary complexity. Qi et al. (2022) used support vector machines, Gaussian mixed models, and Bayesian learning to enhance fluency and vocabulary. Virvou et al. (2012) employed k-means clustering to group students for personalized foreign language instruction, while Hasan and Khaing (2008) used machine learning on an annotated learner corpus to assess proficiency with high accuracy.

These studies demonstrate diverse machine learning applications in PLL, addressing different skills through varied techniques. However, the field of listening skill remain underexplored despite its importance in communicative proficiency. To address this gap, this study proposes Personal Interest-Based Listening-Machine Learning (PiBAL-ML), a system designed to enhance EFL learners' listening skill and assess its effectiveness. Additionally, it explores students' perceptions of the system. Thus, the research questions were formulated as follows:

1. To what extent does the use of PiBAL-ML affect EFL learners' overall listening achievement and performance across different listening task types?
2. How do EFL learners perceive the affordances and challenges of PiBAL-ML as a personalized listening platform?

2. Methodology

2.1. Research design

After designing the PL platform, the current study adopted one of the mixed research approaches, namely explanatory sequential research design (Creswell & Plano Clark, 2018) to investigate the effects of PLL on listening achievement. This study utilized explanatory sequential research as it included an experiment (with one experimental group and one control group) to see the effect of the developed system and then gathered opinions of users to determine their satisfaction/dissatisfaction with this learning experience. The quantitative phase of the study included quasi-experimental design since two groups were assigned to two conditions based on the sections they followed the course. In the qualitative phase, the participants in the experimental group were asked to share their ideas and experiences about the platform developed which they used when they underwent the experimental procedure. The study included both quantitative and qualitative data, with a higher priority on quantitative data since the research initially aimed to explain the effects (Creswell & Plano Clark, 2018) of PLL (QUAN → qual). In accordance with the nature of explanatory sequential design, firstly quantitative data were collected through a listening achievement test developed by the researchers (see Figure 1). The qualitative data, which aimed to understand the participants' views (Creswell & Plano Clark, 2018) on personalized learning and to explain the quantitative data, included open-ended questions given at the end of the intervention to the experimental group. As Creswell and Plano Clark (2011) emphasized, the qualitative data were used to see if qualitative findings helped to explain the quantitative results. For the purpose of the current study, the qualitative data were used to gain a deeper understanding of the participants' views on the effects of the platform. Both data were collected separately and analysed independently. However, the results of both types of data were merged in the interpretation with a discussion of findings based on the relevant literature. Figure 1 illustrates the research design.

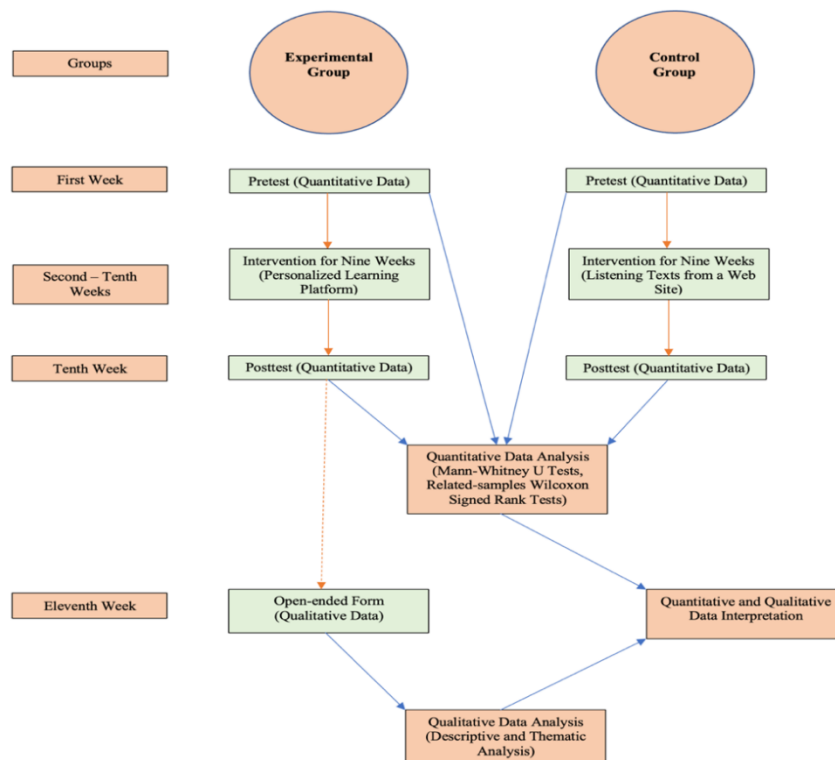


Figure 1. Research design

2.2. Participants

The participants of this study were higher education students who continued their education in English Language Teaching majors learning English as a Foreign Language in a state university in Türkiye (N = 44, n = 20 for the control group, and n = 24 for the experimental group) in the academic year of 2022-2023. Using convenience sampling, Turkish-L1 upper-intermediate to advanced EFL learners from two intact classes participated in the study based on their informed consent. The reason for choosing convenience sampling is that the participants were studying at the university where the researchers worked at. Further, to see the initial differences (Farrokhi & Mahmoudi-Hamidabad, 2012) (e.g. differences between the participants' pretest scores), tests were run to see if the two groups differ in terms of their pretest scores. Two intact groups were assigned to two conditions based on the course sections they were enrolled in. Initially, the study included 60 participants; however, 16 participants were subsequently eliminated, 10 from the control group and 6 from the experimental group, due to their failure to complete one or more of the assigned tasks or tests. For the qualitative phase, only volunteers from the experimental group were included to respond to the open-ended questionnaire.

2.3. Procedure

Personalization in PiBAL-ML was achieved through a dual approach. Firstly, the platform offered listening materials tailored to the individual interests of learners, covering topics like sports, culture, technology, and more. Secondly, a variety of activity types, including fill-in-the-gaps, matching, true/false, and multiple-choice exercises, were integrated for each listening text. The system dynamically adjusted the number of questions within these activities based on learners' success rates. A comprehensive explanation of this procedure is provided across two dimensions: system design and the treatment.

2.3.1. System design

Personalization in PiBAL-ML was achieved through a dual approach. Firstly, the platform offered listening materials tailored to the individual interests of learners, covering topics like sports, culture, technology, and more. Secondly, a variety of activity types, including fill-in-the-gaps, matching, true/false, and multiple-choice exercises, were integrated for each listening text. The system dynamically adjusted the number of questions within these activities based on learners' success rates. A comprehensive explanation of this procedure is provided across two dimensions: system design and the treatment.

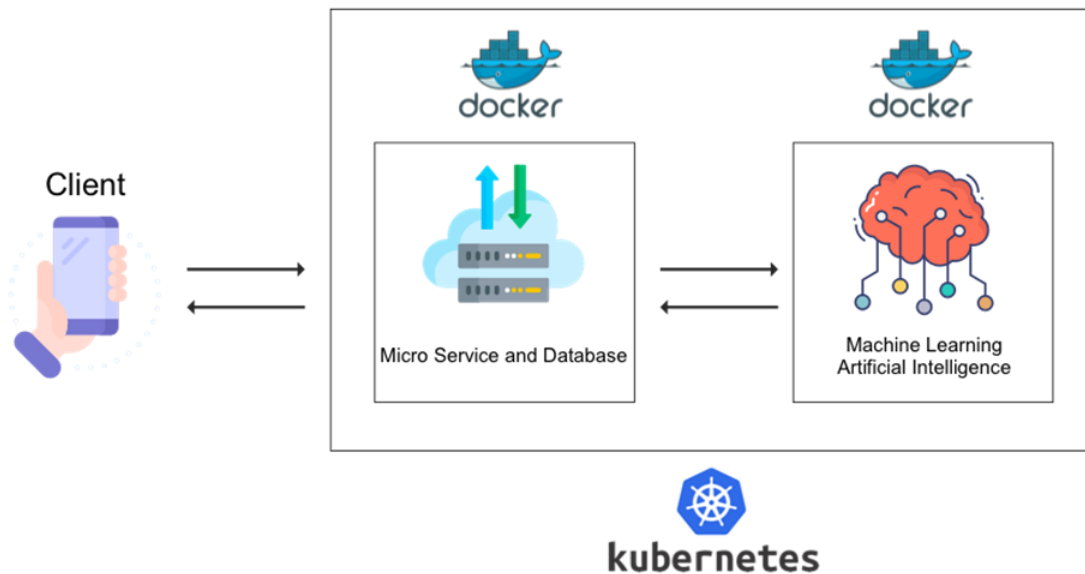


Figure 2. Summary of the system architecture

In the initial phase, the Professional Orientation Inventory questionnaire (Kepçeoğlu, 2019), adapted to 12 questions, was administered to 164 individuals from various fields to identify their daily life interests. The identified fields of interest included technology, business life, health, different cultures, sports, and fine arts.

Subsequently, preprocessing algorithms were applied to the dataset, and 20%, 10%, and 70% of the data were randomly selected to create testing, validation, and training sets, respectively. Among the 12 questions, the last one represented the class (field of interest), while the others were features. Six classes (technology, health, business life, cultures, sports, and fine arts) were determined. The final model phase computed a probability of interest for each class label. K-nearest neighbour, random forest, and artificial neural networks models were trained and tested, with hyperparameters optimized using Bayesian optimization techniques, proving faster and more effective than traditional methods (Jones, 2001; Wu et al., 2019). The random forest model achieved the best results with 78.12% accuracy, 75.38% precision, 65.33% recall, and an 86.32% true negative rate, as detailed in Görmez et al. (2022). This model was used in our system as the area of interest predictor.

Following the machine learning phase, artificial intelligence algorithms were designed and developed. Five scores were used: overall score, fill in the blanks score, multiple-choice score, true-false score, and matching score. The overall score determined the students' English level, which the teacher initialized, with thresholds set for adjusting the level. If the overall score falls below a lowest threshold, the level (upper intermediate, advanced) is lowered and if the overall score rises above a highest threshold, the level is increased. For the number of question types, a vector V with four values, where each of them represents the one question type, is assigned. The number of each activity is computed with the formula of $ni = Vi \times N$. In this formula, N represents the total number of activities, ni represents the number of questions for i th question type and Vi represents the i th value of assigned vector V . The vector values are initialized with 0.25 when the student first enters the system. After each exercise, vector values are updated with the formula of $Vi = (Vi + Si \sum_{k=1}^4 Sk) \div 2$. In this formula, Vi represents the i th value of assigned vector V , Si represents the score for i th question type and $\sum_{k=1}^4 Sk$ represents the total score for all activity types.

In PiBAL-ML, students are first registered in the system by the teacher. Upon their first login, the designed questionnaire is administered, excluding the 12th question, which is not visible to learners. Students can then request content from the system, which the machine learning and artificial intelligence module determines. After completing the exercises related to the content, the artificial intelligence module is updated, and this process is repeated for each new content. Figure 3 illustrates the system's flow diagram.

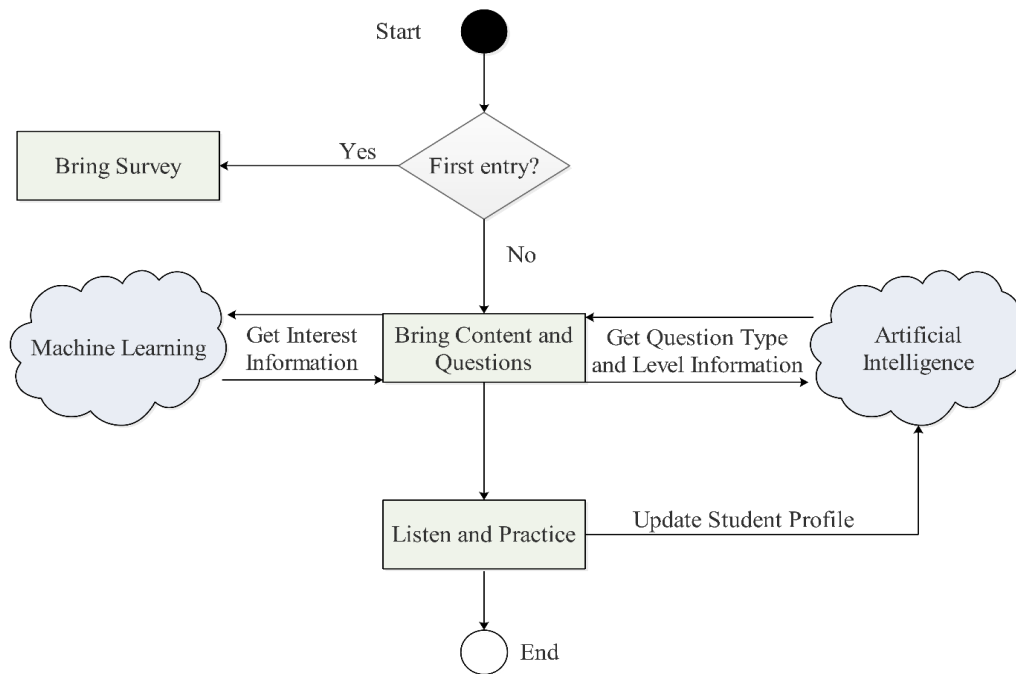


Figure 3. Flow diagram of the proposed system

2.3.2. Treatment

PiBAL-ML was piloted with 10 ELT students for a week, one year prior to the actual treatment, to identify any potential issues. These students were enrolled in the Listening and Pronunciation course, and the pilot was integrated into this course, aligning with the target group. Following this pilot, various issues were identified, such as failure to navigate previous questions, which were subsequently addressed. After resolving these issues, the actual treatment commenced. Before treatment, a listening achievement test was conducted to both groups as pretest. While the experimental group used the proposed system for one semester, the control group also used online listening platform of British Council which was not personalized, and all the students were studying on the same listening at the same time. Both groups received equal exposure in terms of number of texts, difficulty level, time-on-task, task types, and teacher support throughout the treatment period. Listening texts and exercises from the British Council were also incorporated into the content of the system with the necessary permissions obtained prior to use. 20 listening texts and their activities were used by adding different type of activities by the researchers. Participants in the experimental group logged into the system, completed a questionnaire predicting their areas of interest, and engaged in listening activities suggested to them based on these predictions. At the end of semester, both the control and experimental groups underwent the achievement test again as a post-test. Additionally, the experimental group was provided with an open-ended form to express their opinions on the system.

2.4. Instruments

Qualitative and quantitative data were gathered through two instruments. Firstly, a listening achievement test developed by the researchers was used as the pretest and post-test. The Cronbach alpha of this test was found to be .69 when participants' post-test scores were used. The tests consisted of four question types including: a) multiple-choice questions, b) true-false questions, c) fill in the blanks questions, and d) matching questions. The participants from both groups listened to a text at the C1 level, as determined and designated by the national language exam, and subsequently responded to comprehension questions. The second instrument used in the study was an open-ended form developed by the researchers. It consisted of three questions. The first question asked students to rate the platform out of five, the second one was directed to the participants to see what they thought on the usefulness of the platform, and the final question asked participants to write down what they thought was negative about the platform. The form was conducted to the experimental group to gain a deeper understanding of the effects of the platform proposed.

2.5. Data analysis

The quantitative data were analysed through SPSS 26. Since the data violated the normality assumption, 10 Mann-Whitney U tests were used when comparing the participants' pre and post-test scores. The reason for using 10 tests is that the listening achievement test consisted of four sub-items, meaning that eight tests were used to compare all the sub-components of the tests, and two were used to compare the participants' overall scores. In addition, to see the participants' improvements from pretest to post-test scores, 10 related-samples Wilcoxon signed rank tests were used.

For the analysis of qualitative data and respond to the second research question, the research adopted qualitative thematic analysis method (Braun & Clarke, 2006) using NVivo, and descriptive statistics using SPSS 26. Thematic analysis involved two approaches: a) deductive, and b) inductive coding. For the deductive coding phase, the questions of the open-ended questionnaire formed two main themes: a) affordances, and b) challenges. Under these two predetermined codes, several themes occurred as a result of the inductive coding phase. Before starting the coding phase, the researchers read the data several times. Later, they categorized the participants' answers to the questions under two predetermined themes. After categorizing these answers, they read the data under predetermined themes several times to establish different themes using inductive coding. For this phase, two researchers read the data separately and came together to discuss their results. Inter-coder reliability was calculated as .95 using MAXQDA 2018. Conflicts over themes were resolved through discussions between the two researchers who gave courses to the participants.

2.6. Ethical considerations

Conducting research within the scope of a course brings about some ethical issues. To avoid these issues, Murray et al. (1996) highlight that the study should intend to support the learners in the learning process. In the case of the current study, with the PiBAL-ML, the study aimed to improve the learners' listening skills. Since the study was conducted within the scope of Listening and Pronunciation course, the study was well-suited for the objectives of the course. In addition, the study followed the Scientific Research and Publication Ethics Directive published by the Council of Higher Education in Türkiye. Further, the researchers applied for ethical permission. Ethical approval for the study was obtained from Sivas Cumhuriyet University Scientific Research and Publication Ethics Social and Human Sciences Board dated 02 February 2021 and numbered E-60263016-050.06.04-9049.

3. Results

The results and findings were presented in accordance with two research questions. The answer of first research question yielded quantitative results, and the qualitative data were gathered to answer the second research question.

3.1. Quantitative results

3.1.1. Between-group comparisons

Five Mann-Whitney U tests were conducted to compare the pretest scores. The results of the tests indicated no significant differences between the participants' pretest scores in multiple-choice questions ($U = 310.00$, $p = .06$), true-false questions ($U = 215.00$, $p = .48$), fill in the blank questions ($U = 183.50$, $p = .15$), matching questions ($U = 237.00$, $p = .94$), and overall scores ($U = 232.50$, $p = .86$) (see Table 1).

Table 1: *The descriptive statistics of participants regarding their pre-test scores*

Test type	Experimental group mean (SD)	Control group mean (SD)
Multiple-choice	4.17 (.92)	4.65 (.59)
True-false	4.63 (.65)	4.50 (.69)
Fill in the blanks	3.96 (.86)	3.50 (1.00)
Matching	4.38 (.77)	4.30 (.92)
Total	17.13 (1.57)	16.95 (2.09)

Five more Mann-Whitney U tests were run to compare the participants' post-test scores. The results highlighted no significant differences between the participants' scores in multiple-choice questions ($U = 177.00$, $p = .08$), true-false questions ($U = 214.00$, $p = .42$), fill in the blank questions ($U = 169.00$, $p = .08$), matching questions ($U = 221.00$, $p = .58$), and overall scores ($U = 172.00$, $p = .10$) (see Table 2).

Table 2: *The descriptive statistics of participants regarding their post-test scores*

Test type	Experimental group mean (SD)	Control group mean (SD)
Multiple-choice	4.75 (.44)	4.20 (1.11)
True-false	4.71 (.62)	4.55 (.83)
Fill in the blanks	4.13 (.90)	3.55 (1.05)
Matching	4.46 (1.18)	4.45 (.89)
Total	18.04 (1.81)	16.75 (2.75)

3.1.2. Within-group comparisons

3.1.2.1. Control group

Five Wilcoxon-signed rank tests were run to see whether the participants in the control group developed their listening skills. The results indicated no significant differences for true-false questions ($z = -.26$, $p = .79$), fill in the blanks questions ($z = -.24$, $p = .81$), matching questions ($z = -.54$, $p = .59$), and overall scores ($z = -.48$, $p = .63$). However, the results indicated a significant difference in the participants' scores in multiple-choice questions, $z = -2.08$, $p = .04$, with a moderate effect size, $r = .33$.

3.1.2.2. Experimental group

Five more Wilcoxon-signed rank tests were conducted to see whether the participants in the experimental group developed their listening skills. The results indicated no significant differences for true-false questions ($z = -.49$, $p = .63$), fill in the blanks questions ($z = -.88$, $p = .38$), matching questions ($z = -.59$, $p = .55$), and overall scores ($z = -1.81$, $p = .07$). However, the results indicated a significant

difference in the participants' scores in multiple-choice questions, $z = -2.43$, $p = .02$, with a moderate effect size, $r = .35$.

3.2. Qualitative findings

3.2.1. Rating findings

The first question of the questionnaire was directed to participants to rate the platform out of five. The results indicated that the participants overall had a positive attitude towards the system ($M = 4.00$, $SD = .66$).

3.2.2. Themes

The responses from 15 participants, collected in response to three open-ended questions, underwent a thematic analysis. This analysis yielded two overarching themes: a) affordances b) challenges. Within these two primary themes, a total of six subcategories emerged (two for challenges and four for affordances). Consequently, the findings culminated in the identification of five distinct overarching themes (see Table 3 and Figure 4). Within affordances, design of the system category included two different themes: feedback and user-friendliness, and for challenges, the same category included three different themes: evaluation, technical problems, and other.

Table 3: *The frequency of the participants' views*

Themes	f	%
Affordances	29	64.44
Design of the system	5	11.11
Feedback	2	4.44
User-friendliness	3	6.67
Activities	5	11.11
Improvement of listening skill	5	11.11
Personalized learning	14	31.11
Challenges	16	35.56
Design of the system	14	31.13
Evaluation	3	6.67
Technical problems	4	8.90
Other	7	15.56
Activities	2	4.44

3.2.2.1. Affordances

The responses provided by the participants in the questionnaire were analysed to discern emerging themes within the predefined category of "affordances." The participants' favourable views regarding PiBAL-ML consisted of four key domains: the system's design, the variety of activities, enhancements in listening skills, and, notably, personalized learning, which appeared to be of primary importance to them.

The first theme emerged under affordances focused on the system design with two sub-themes, namely feedback and user-friendliness. Under the feedback theme, we see reflections on the system's immediate feedback design. The system helped the participants to see their mistakes. In fact, one of the participants stated that "[The system] made me see my shortcomings" (P. 5). The other sub-theme, user-friendliness, focused on the listening materials' quality, and the design of the interface. One participant stated that "Videos and questions were in good quality" (P. 11), and another participant stated that "It is fast and useful" (P. 9). Under the activities theme, the participants' responses were related to the types

of questions such as multiple-choice questions, different activities, and the quality of listening activities. One of the participants stated that *“We can improve ourselves with different exercises”* (P. 1). The improvement of listening skill theme included the participants’ reflections on their views of the platform’s effectiveness on listening skill. One of the participants wrote that *“They are beneficial in terms of listening skills”* (P. 10). The last theme under the affordances, personalized learning, which received the highest number of reflections, focused on the participants’ views on the effectiveness of personalized learning, which yielded positive feedback. One of the participants stated that *“Doing activities according to my interests allowed me to use the platform more effectively and carefully”* (P. 15).

3.2.2.2. Challenges

The responses of participants to the questionnaire were analysed to determine themes for the second predetermined theme: challenges. As a result of the qualitative analysis, two categories emerged: a) design of the system, and b) activities.

The first theme under challenges consisted of three categories focusing on evaluation, technical problems, and other issues. The evaluation theme usually focused on some mistakes regarding the participants’ overall scores rather than giving the wrong feedback. For example, one participant stated that *“It shows wrong score at the end”* (P. 10). Technical problems usually consisted of sound issues and access to internet issues. One participant stated that *“Some errors such as downloading problems about listening files make it hard to concentrate”* (P. 3). The other theme included the participants’ negative views on the design of the system such as not being able to go back to the missed point, lack of content in terms of number, modality of the input, etc. One participant’ statement *“not being able to watch the video while listening”* (P. 5) was an issue related to the modality of the input. The second theme, activities, focused on the difficulty and the length of the exercises. While one participant, focusing on the length, stated that *“Some exercises were too long or too short”* (P. 14), other stated that *“Some of the exercises are a little bit hard on the listening levels”* (P. 2).

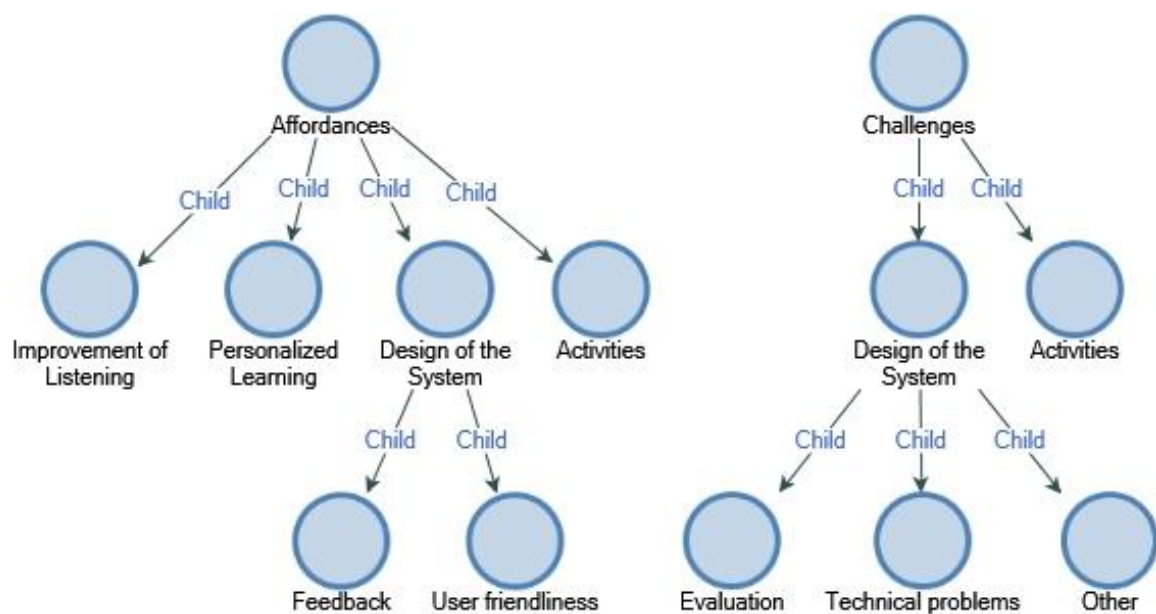


Figure 4. A project map of themes

4. Discussion

The main purpose of this paper was to examine the effect of PiBAL-ML on the listening skills of EFL learners, their satisfaction with the platform, and their opinions on personalized listening learning (PLL). The quantitative analysis yielded several important results. PiBAL-ML utilized various types of content and listening activities to provide personalization. Using a single type of activity would contradict the nature of personalized learning, so four different types of activities, fill in the blanks, matching, multiple choice, and true-false, were employed both in the system and in the achievement test. While the overall scores and other types of activities did not show significant differences between groups, the success in multiple-choice questions was significant in favour of the experimental group. This result can be commented through different lenses. Firstly, the system itself was effective for improving listening skills of learners. Secondly, the multiple-choice questions were regarded as more favourable, familiar or easy by students as control group was successful in pretest and experimental group was successful in post-test. Turkish education system has some essential exams for entering high school and university and getting job and these essential exams use only multiple-choice questions. Multiple-choice tests are one of the most common and important assessment types in Türkiye, so the Turkish students are more accustomed to multiple choice tests than any other measuring and evaluation techniques.

The quantitative results of this mixed-method study showed no significant differences between the control and experimental groups in both pretest and post-test results, regardless of the test type. The overall scores suggest that personalization based on personal fields of interest (e.g., sports, fine arts, technology) did not significantly impact the improvement of learners' listening skills. Although there is not much finding related to the effect of PLL on listening skills, the experimental studies focusing on different skills contradict with this study as they generally show positive results on behalf of experimental groups using personalized systems (Chen & Hsu, 2008; Hsu et al., 2013; Kwon et al., 2010; Zou, 2011). However, this finding should be read with caution. Firstly, the previous studies focus on different skills and the duration of improving these skills may show variety. For example, learning new vocabulary may require less time from improving communicative skills. Secondly, most of the studies in literature did not include control group in their study (e.g., Chen & Hsu, 2008; Kwon et al., 2010; Zou, 2011) but concluded with the positive results by analysing pre-test and post-test of the experimental group. This study also found that the experimental group got higher score in post-test than pre-test. Moreover, previous studies used different techniques for personalization while this study benefited from machine learning. Lastly, as mentioned before, acquiring communicative skills such as listening and speaking requires more time. The system was implemented one semester in this study. Thus, a longer treatment may lead to different results.

The participants' ratings regarding the platform indicated that they found the platform quite useful, which showed that learners' satisfaction was high. Moreover, qualitative results indicated that the participants' responses to the open-ended questions mostly focused on the positive sides (64.44 %) of the system proposed compared to the negative sides (35.56 %). The learners regarded the system as user-friendly and stated that the system presented different activities and instant feedback, which improved their listening skills. This finding is parallel with the previous research. For example, the teachers regarded the personalized system useful for language learning in the study of Petersen et al. (2009). Similarly, learners believed that personalized reading ability system improved their success in reading (Hsu et al., 2013). In contrast to benefits of PiBAL-ML, some challenges were addressed by the participants. While one student found the listening texts difficult, others regarded them as either too short or too long. It should be noted, however, that these issues were not related to the system itself, but

rather to the content sourced from the British Council website. To better meet learner needs and expectations, greater care could be taken in the content selection process in future implementations. The responses of the experimental group participants who took part in the questionnaire highlighted the importance of multiple-choice questions, correlating with the quantitative findings. For instance, one participant stated, "multiple-choice questions were helpful for me" (P. 12). Therefore, the improvements of the experimental group in multiple-choice questions can be attributed to the intervention rather than being coincidental.

This study contributes to the growing body of research on personalized and technology-enhanced language learning by examining the effects of PiBAL-ML, a machine learning-based personalized listening platform, on EFL learners' listening skills and satisfaction. While statistically significant differences between the control and experimental groups were limited to multiple-choice question performance, the overall findings suggest that PiBAL-ML holds promise as a learner-centered instructional tool. The high levels of learner satisfaction and the predominantly positive qualitative feedback indicate that the system was perceived as beneficial by the participants. These findings underscore the importance of designing language learning systems that not only target skill development but also prioritize learner engagement, usability, and motivation. As personalized learning technologies continue to evolve, future research should explore ways to refine content selection, extend implementation periods, and leverage the full capabilities of machine learning to better support the diverse needs of EFL learners.

5. Conclusions

Despite its preliminary nature, this study concluded with a positive role of machine learning based PLL in English learning. Even if the students in experimental group did not outperform control group in overall achievement, they were found to be more successful in multiple choice test type which is commonly used in Türkiye. Moreover, the learners' opinions regarding the use of personalized system were generally positive. These findings suggest utilizing PL in language education, as it has benefits such as providing activities and materials based on learners' own interest and needs. In the light of findings, it can be proposed that teachers should be more aware of PLL advantages and their students' individual differences. Besides, the policy makers and administrators should help teachers to create an atmosphere where PLL is adopted effectively. More effort can be given to adjust the curriculum and teaching materials to learners' needs, speed, wishes and lacks.

This study has some limitations including small sample size and short-term treatment which lasted for one semester. It is known that improving a linguistic skill requires time. Moreover, utilizing technology is not lack of limitations and some technical problems occur. The participants of this study also mentioned about these problems which can affect the results. In the light of these limitations, this study can be replicated with a longer treatment duration and by minimizing the technical problems. Moreover, six fields of interest such as sports, business life, culture, fine arts, technology, and health were included in the system content and four type of activities such as fill in the gap, matching, multiple choice and true/false. More activities and fields of interest can be adopted as individuals have various interests and needs. Lastly, different skills and learners with different English level may be searched in further research within machine-based learning PLL concept as the literature remains limited, which means that there is much to explore.

6. Statements

6. 1. Conflict of Interest

There is no conflict of interest to declare.

6. 2. Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

6. 3. Funding

There is no funding information to declare.

6. 4. Acknowledgement

We thank Detaysoft Technology Company for their support.

6. 5. Ethical Approval

The study followed the Scientific Research and Publication Ethics Directive published by the Council of Higher Education in Türkiye. Further, the researchers applied for ethical permission. Ethical approval for the study was obtained from Sivas Cumhuriyet University Scientific Research and Publication Ethics Social and Human Sciences Board dated 02 February 2021 and numbered E-60263016-050.06.04-9049.

6. 6. Authors Contribution

Kübra OKUMUŞ DAĞDELER: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization

Yasin GÖRMEZ: Software, Writing – review & editing, Writing – original draft, Resources, Conceptualization

Samet Çağrı KIZKAPAN: Visualization, Formal analysis, Data curation, Methodology, Writing – original draft

References

- Adams Becker, S., Cummins, M., Davis, A., Freeman, A., Hall Giesinger, C., & Ananthanarayanan, V. (2017). *NMC Horizon Report: 2017 Higher Education Edition*. The New Media Consortium.
- Basham, J. D., Hall, T. E., Carter, R. A., & Stahl, W. M. (2016). An operationalized understanding of personalized learning. *Journal of Special Education Technology*, 31(3), 126-136. <https://doi.org/10.1177/0162643416660835>.
- Bernal Parraga, A. P., Coronel Ramírez, E. A., Aldas Macias, K. J., Carvajal Madrid, C. A., Valarezo Espinoza, B. D. C., Vera Alcivar, J. G., & Chávez Cedeño, J. U. (2025). The Impact of Artificial Intelligence on Personalized Learning in English Language Education. *Ciencia Latina Revista Científica Multidisciplinar*, 9 (1), 5500-5518. https://doi.org/10.37811/cl_rcm.v9i1.16234
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>.
- Bulger, M. (2016). Personalized learning: The conversations we're not having. *Data and Society*, 22(1), 1-29.
- Chen, C.-M., & Chung, C.-J. (2008). Personalized mobile English vocabulary learning system based on item response theory and learning memory cycle. *Computers & Education*, 51(2), 624-645. <https://doi.org/10.1016/j.compedu.2007.06.011>

- Chen, C.-M., & Hsu, S.-H. (2008). Personalized intelligent mobile learning system for supporting effective English learning. *Educational Technology & Society*, 11(3), 153-180.
- Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning: Topic modelling and knowledge mapping. *Educational Technology & Society*, 24(1), 205-222.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.
- Fang, L., Tuan, L. A., Hui, S. C., & Wu, L. (2017). Personalized question recommendation for English grammar learning. *Expert Systems*, 35(2). <https://doi.org/10.1111/exsy.12244>
- Farrokhi, F., & Mahmoudi-Hamidabad, A. (2012). Rethinking convenience sampling: Defining quality criteria. *Theory and Practice in Language Studies*, 2(4). <https://doi.org/10.4304/tpls.2.4.784-792>
- Görmez, Y., Okumuş Dağdeler, K., & Kavuklu, M. (2022). Kişiselleştirilmiş yabancı dil öğrenimi için makine öğrenmesi yöntemleriyle ilgi alanı tahmini. *Yükseköğretim ve Bilim Dergisi*, 12(1), 111-121. <https://doi.org/10.5961/higheredusci.982740>
- Griol, D., Baena, I., Molina, J. M., & de Miguel, A. S. (2014). A multimodal conversational agent for personalized language learning. In C. Ramos, P. Novais, C. E. Nihan, & J. M. Corchado Rodríguez (Eds.), *Ambient Intelligence—Software and Applications* (pp. 13-21). Springer International Publishing. https://doi.org/10.1007/978-3-319-00566-9_2
- Hasan, M. M., & Khaing, H. O. (2008). Learner corpus and its application to automatic level checking using machine learning algorithms. In *Proceedings of the 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (Vol. 1, pp. 25–28). <https://doi.org/10.1109/ECTICON.2008.4600364>
- Hsu, C.-K., Hwang, G.-J., & Chang, C.-K. (2013). A personalized recommendation-based mobile learning approach to improving the reading performance of EFL students. *Computers & Education*, 63, 327-336. <https://doi.org/10.1016/j.compedu.2012.12.004>
- Jiang, H. (2025). Deep learning based personalized English listening learning path recommendation algorithm. *Systems and Soft Computing*, 7. <https://doi.org/10.1016/j.sasc.2025.200210>
- Jones, D. R. (2001). A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21, 345–383. <https://doi.org/10.1023/A:1012771025575>
- Jung, J., & Graf, S. (2008). An approach for personalized web-based vocabulary learning through word association games. In *International Symposium on Applications and the Internet* (pp. 325-328). IEEE. <https://doi.org/10.1109/SAINT.2008.95>
- Kepçeoğlu, M. (2019). *Eduolog.com - Özgün ve Bağımsız Çevrimiçi Eğitim Rehberi*. Accessed on <https://www.eduolog.com/>
- Kerr, P. (2016). Adaptive learning. *ELT Journal*, 70(1), 88-93. <https://doi.org/10.1093/elt/ccv055>
- Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271–289. <https://doi.org/10.1017/S0958344008000335>
- Kwon, D., Lim, H., Lee, W., Kim, H., Jung, S., Suh, T., & Nam, K. (2010). A personalized English vocabulary learning system based on cognitive abilities related to foreign language proficiency. *Transactions on Internet and Information Systems*, 4(4), 595–617.
- Looi, C.-K., Wong, L.-H., So, H.-Y., Seow, P., Toh, Y., Chen, W., Zhang, B., Norris, C., & Soloway, E. (2009). Anatomy of a mobilized lesson: Learning my way. *Computers & Education*, 53, 1120-1132. <https://doi.org/10.1016/j.compedu.2009.05.019>
- Murray, H., Gillese, E., Lennon, M., Mercer, P., & Robinson, M. (1996). *Ethical principles in university teaching*. Accessed on <https://files.eric.ed.gov/fulltext/ED407899.pdf>
- Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). *Continued progress: Promising evidence on personalized learning*. RAND Corporation.

- Peng, H., Ma, S., & Spector, J. M. (2019). Personalized adaptive learning: An emerging pedagogical approach enabled by a smart learning environment. *Smart Learning Environments*, 6(1). <https://doi.org/10.1186/s40561-019-0089-y>
- Petersen, S. A., Markiewicz, J.-K., & Bjørnebekk, S. S. (2009). Personalized and contextualized language learning: Choose when, where and what. *Research and Practice in Technology Enhanced Learning*, 4(1), 33-60. <https://doi.org/10.1142/S1793206809000635>
- Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, 7(3), 135–153. <https://doi.org/10.2478/jolace-2019-0025>
- Qi, S., Liu, L., Kumar, B. S., & Prathik, A. (2022). An English teaching quality evaluation model based on Gaussian process machine learning. *Expert Systems*, 39(6). <https://doi.org/10.1111/exsy.12861>
- Ross, M., Graves, C. A., Campbell, J. W., & Kim, J. H. (2013). Using support vector machines to classify student attentiveness for the development of personalized learning systems. *12th International Conference on Machine Learning and Applications*. <https://doi.org/10.1109/ICMLA.2013.66>
- Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025). The role of large language models in personalized learning: a systematic review of educational impact. *Discover Sustainability*, 6(243). <https://doi.org/10.1007/s43621-025-01094-z>
- Traxler, J. (2009). *Current state of mobile learning: Transforming the delivery of education and training*. Accessed on <http://www.aupress.ca/index.php/books/120155>
- Troussas, C., Chrysafiadi, K., & Virvou, M. (2018). Machine learning and fuzzy logic techniques for personalized tutoring of foreign languages. In C. Penstein Rosé et al. (Eds.), *Artificial Intelligence in Education* (pp. 358-362). Springer International Publishing. https://doi.org/10.1007/978-3-319-93846-2_67
- Troussas, C., Chrysafiadi, K., & Virvou, M. (2019). An intelligent adaptive fuzzy-based inference system for computer-assisted language learning. *Expert Systems with Applications*, 127, 85–96. <https://doi.org/10.1016/j.eswa.2019.03.003>
- Troussas, C., Virvou, M., & Espinosa, K. J. (2015). Using visualization algorithms for discovering patterns in groups of users for tutoring multiple languages through social networking. *Journal of Networks*, 10(12), 668-674. <https://doi.org/10.4304/jnw.10.12.668-674>
- U.S. Department of Education, Office of Educational Technology. (2017). *Reimagining the role of technology in education: 2017 National Education Plan update*. U.S. Department of Education. Washington, DC
- Virvou, M., Troussas, C., & Alepis, E. (2012). Machine learning for user modelling in a multilingual learning system. In *International Conference on Information Society* (pp. 292-297).
- Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology*, 105(4), 932–945.
- Wang, Z. (2023). Online language education recommendation based on personalized learning and edge computing. *Internet Technology Letters*, 408. <https://doi.org/10.1002/itl2.408>
- Wu, J., Chen, X.-Y., Zhang, H., Xiong, D.-L., Lei, H., & Deng, S.-H. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. *Journal of Electronic Science and Technology*, 17(1), 26–40. <https://doi.org/10.11989/JEST.1674-862X.80904120>
- Wu, T.-T., Huang, Y.-M., Chao, H.-C., & Park, J. H. (2014). Personalized English reading sequencing based on learning portfolio analysis. *Information Sciences*, 257, 248-263. <https://doi.org/10.1016/j.ins.2011.07.021>
- Xu, Y., Zhang, M., & Gao, Z. (2020). The construction of distance education personalized learning platform based on educational data mining. In J. Abawajy, K.K. Choo, R. Islam, Z. Xu, & M. Atiquzzaman (Eds.), *ATCI 2019. Advances in Intelligent Systems and Computing*, 1017. Springer, Cham. https://doi.org/10.1007/978-3-030-25128-4_134
- Yang, C. C. Y., & Ogata, H. (2023). Personalized review learning approach for improving behavioural engagement and academic achievement in language learning through e-books. *Education and Information Technologies*, 28, 1491–1508. <https://doi.org/10.1007/s10639-022-11245-8>

- Yeung, C. Y., & Lee, J. (2018). Personalized text retrieval for learners of Chinese as a foreign language. *Proceedings of the 27th International Conference on Computational Linguistics*, 3448-3455. Accessed on <https://aclanthology.org/C18-1292>
- Zhang, L. (2021). A new machine learning framework for effective evaluation of English education. *International Journal of Emerging Technologies in Learning*, 16(12), 142–154. <https://doi.org/10.3991/ijet.v16i12.23323>
- Zou, Y. (2011, December). Personalized automatic recommendations for the web-based autonomous language learning system based on data mining technology. In *IEEE International Symposium on IT in Medicine and Education* (Vol. 2, pp. 326-329). IEEE.



© This study is submitted for submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>)