ISSN 2411-9563 (Print) European Journal of Social Science March 2026
ISSN 2312-8429 (Online) Education and Research Volume 13, Issue 1

Submitted: 09/01/2026 - Accepted: 11/02/2026 - Published: 29/03/2026

Adaptive Learning Platforms in English Education: Enhancing Comprehensive Skills and Social Interaction in the Digital Classroom

Lizhen Liu[1*]

[1*]*School of Foreign Languages,Jining Normal University,*
Inner Mongolia Autonomous Region， Ulanqab， China
jenny202511@126.com

DOI: 10.26417/xscbj161

Abstract:

Advances in computer-assisted- language learning (CALL) promise personalized instruction, yet most adaptive systems focus on isolated skills or vocabulary. This study reports a 12-week quasi-experimental mixed-methods evaluation of an adaptive platform that integrates the Vygotsky zone of proximal development (ZPD), Bayesian Knowledge Tracing (BKT), and Item Response Theory (IRT) to foster comprehensive proficiency across listening, speaking, reading, and writing. ANCOVA controlling for pre-test scores indicated significant gains in all skills for the adaptive group ($F = 19.30–44.65$, $p \leq 1.82 \times 10^{-5}$); effect sizes ranged from moderate for writing ($d = 0.62$) too large for listening, speaking, and reading ($d = 0.84–0.93$). Learners reported higher perceived ease of use, usefulness, and enjoyment, and these constructs correlated with proficiency gains. Qualitative data revealed enhanced motivation, autonomy, and reduced anxiety. Framing the design as a controlled short-term- intervention, the study highlights the platform's pedagogical affordances and psychometric calibration while acknowledging the limitations of duration, sample homogeneity, and anxiety measurement. These findings support the feasibility of integrating cognitive scaffolding with latent mastery estimation to deliver balanced, adaptive English instruction.

Keywords: Digital Pedagogy, Educational Technology, Adaptive Learning, Language Education, Student Engagement

1. Introduction

Computer assisted language learning has evolved from fixed drill and practice programs and has taken the shape of the intelligent computer system that tailors the instruction on-the-fly (Tanaka et al., 2025). The intersection of artificial intelligence, natural language processing, and automatic speech recognition has created intelligent CALL platforms that adjust task difficulty and feedback to the current proficiency of the learner and implement the concept of the zone of proximal development by Vygotsky and dynamically regulate cognitive load (Meini et al., 2025). In intelligent tutoring systems, the meta analyses are said to have yielded gains of about half a standard deviation compared to traditional teaching, and the adaptive algorithms like Bayesian Knowledge Tracing and Item Response Theory have facilitated real-time sequencing and a psychometrically sound assessment. This has changed how language is taught and now, personalization at scale can be achieved (Wang & Guo, 2025). Consequently, effective curricula need to incite and inculcate each of the skills individually. However, the existing literature on intelligent CALL is very specific to a particular aspect (vocabulary learning or reading comprehension) and it is not yet clear whether it is possible to use adaptive systems to encourage integrated development in all four modalities (Liu, 2024).

This is complicated by the psychology of learners. Learning-learner agency is one of the key factors contributing to the acquisition of the language, as it entails the ability to formulate goals, make decisions, and manage yourself. Adaptive platforms have the potential to promote agency, which can enable learners to select topics, change pacing, and track their progress, yet can also increase foreign language anxiety or cognitive load in case their algorithms are opaque (Huang et al., 2024). The Technology Acceptance Model states that perceptions of usefulness and ease of use are the determinants of educational technology adoption adopted by learners. It is imperative to understand these affective and perceptual dimensions to maximize the benefits and prevent the negative effects that are not intended (Herzallah & Makaldy, 2025).

Considering the high rate of the development of the AI-mediated platforms, some gaps remain. The combined study of listening, speaking, reading, and writing is not often explored empirically (Sari, 2023). Very little experiments have investigated how adaptive algorithms based on Bayesian Knowledge Tracing and Item Response Theory can be used to affect multiple modalities at the same time, or how adaptive learning can be compared to non-adaptive e learning (Rashid et al., 2025). Furthermore, the interplay among affective variables, i. e. motivation, agency, and anxiety, and the results of language proficiency is yet to be studied (Apoki et al., 2022). Lastly, in traditional CALL systems the sequence of tasks offered is one size fits all without taking into account the prior knowledge of the learners and their personal speed and thus frustration, disengagement and inefficient study may ensue (Raj & Renumol, 2022).

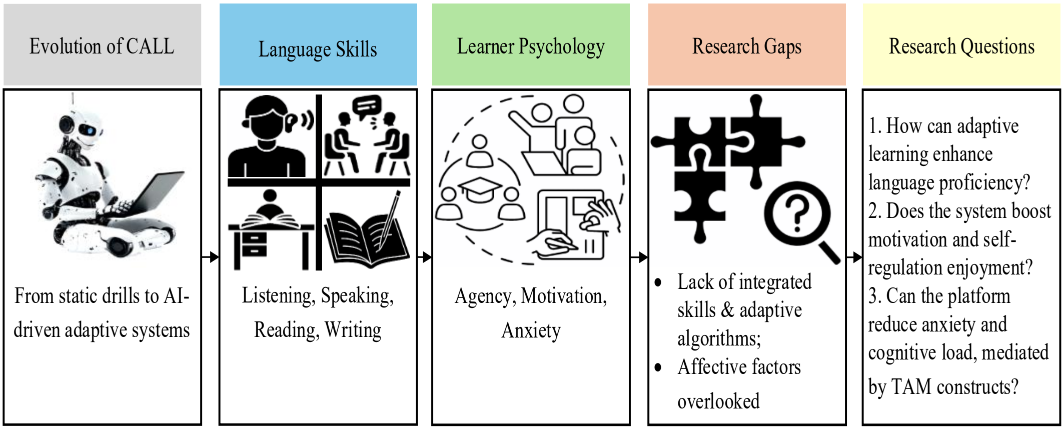


Figure 1. Roadmap of study

This study will address these gaps using a quasi-experimental mixed methods investigation of an adaptive learning platform that will aid comprehensive English proficiency. Overall structure of study is presented in **Figure 1**. Our research questions are as follows: (1) To what extent does the adaptive platform improve the level of listening, speaking, reading and writing compared to the conventional instruction? (2) What is the effect of the platform on learner motivation, self-regulation and perceived enjoyment and which adaptive features mediate the effect? (3) Does the platform lower foreign language anxiety and perceived cognitive load and instill conventional teaching? Which are the mediating effects of perceptions of ease of use and usefulness?

2. Theoretical Framework and Literature Review

Socio cultural and cognitive theories and in particular the Vygotsky perspective of learning as a mediated process with the ZPD as the boundary between tasks that the learner can undertake without any assistance and those that he or she can undertake with assistance, underlie the development of CALL as moving beyond the drill-and-practice approaches to intelligent, adaptive systems (referred to as intelligent tutoring) (Aryadoust et al., 2024). The ZPD instruction is based on scaffolding, modeling, prompting, and guidance by MKO, and through progressive withdrawal as the learner becomes more proficient, and through adaptive technology repeats this same process by constantly checking progress and making hints or corrective feedback as the cognitive capacity allows. Although the Input Hypothesis by Krashen posits that comprehensible input just slightly above the current level of a learner (i+1) is enough to acquire language. Constructivism focuses on the importance of the learners in the active construction of their knowledge based on the experiences and social interactions (Saleem et al., 2021). The adaptive learning systems offer an interactive learning experience in which learners are exposed to customized tasks, which are in tune with their cognitive growth, hence gaining deeper understanding and building of knowledge.

Cognitive Load Theory (CLT) also focuses on the idea that academic instruction should strike the balance between the types of cognitive loads intrinsic, extraneous and germane to avoid overload, and AI-based adaptive reading tools have been shown to vary task difficulty and to give multimodal explanation to promote understanding and reduce anxiety (Su et al., 2021). According to Zeidner and Stoeger (2019), the Self-Regulated Learning (SRL) theory argues that the capacity of learners to monitor and regulate their learning strategies can be very instrumental in academic achievement. The adaptive learning platform has the potential to support the development of self-regulation by allowing learners to achieve specific goals, track their progress, and change their learning strategies in real time through the provision of feedback and custom personalized learning pathways.

Bayesian Knowledge Tracing (BKT) and Item Response Theory (IRT), knowledge-tracing models, are used to estimate the skill mastery of learners over time to support individualized practice, with loss of continuous data: BKT models mastery as a binary latent variable with parameters of initial knowledge, learning rate, slip probability, and guess probability, and the discrete nature of the BKT structure restricts the use of continuous behavioral data, but extended deep learning models like Deep Knowledge Tracing (DKT) can learn complex patterns of mastery with lower interpretability (Bulut et al., 2023). We combine BKT and IRT in the adaptive platform to predict content difficulty and mastery of the learners and create transparency between instructors and learners. This model allows dynamically sequencing tasks to maximize engagement, adaptive testing and real-time ability estimation, and in combination with BKT-based hints and item selection, the learner gains more control and motivation (Yu & Douglas, 2023). Conventional IRT, however, is based on unidimensional ability, which might not be as multidimensional as language proficiency, a problem that has been addressed in recent multidimensional IRT and machine-learning improvements.

3. Literature Review

Both systematic reviews and empirical studies published within a period of 2020 to 2025 support the notion that AI-supported language learning systems can significantly improve engagement and learning performance between students by providing real-time feedback, personalization, and adaptive teaching content delivery and also expose the inability to overcome methodological and ethical issues (Aryadoust et al., 2024). Equally, the analyses of intelligent tutoring systems (ITS) in K-12 settings indicate that learning outcomes are notably contingent on the nature of learners and duration of exposure, and that more time is necessary to counteract the effect of novelty, which supports the perception that AI ought to be used to supplement teachers, but not to replace them (Xu et al., 2019). The reviews on virtual reality (VR), augmented reality (AR), and adaptive learning technologies (ALT) also confirm improvements in vocabulary retention, motivation, and engagement with contextual and immersive experiences, overemphasizing on vocabulary, lack of crosscontext validation, and broader evaluations of full language proficiency, teacher training, and accessibility (Weng & Sing, 2018).

Recent international research further emphasizes that adaptive learning systems can promote social inclusion by supporting learners with diverse linguistic backgrounds, disabilities, and unequal prior educational exposure through flexible pacing and personalized scaffolding (Holmes et al., 2019). Large-scale reviews and policy-oriented studies report that adaptive digital platforms are particularly effective in reducing achievement gaps when aligned with inclusive pedagogical principles rather than purely performance-driven optimization (Gligorea et al., 2023; van der Vlies, 2020). Empirical evidence from higher education contexts also suggests that adaptive learning environments enhance participation and persistence among marginalized learners by fostering autonomy, self-regulation, and perceived instructional fairness (Mirata et al., 2020).

Wider surveys of adaptive AI/ML-based learning systems demonstrate clear advantages, such as personalized learning experiences, dynamically suggested recommendations, automatic updating of learner models, and identifying knowledge gaps, and live feedback and personalized challenges enhance engagement and learning effectiveness, but offer challenges, including cold-start problems, the complexity of the system, privacy concerns, and excessive reliance on the technology, which demonstrates the value of explainable and context-sensitive algorithms to support trust and transparency (Angelopoulos et al., 2019). These trends can be found in the empirical findings across the language skills: a mixed-methods study indicated significant gains on reading comprehension, motivation, and reduction of anxiety with the help of an AI-enhanced reading platform (Wang & Guo, 2025). Similar studies on listening and pronunciation demonstrate that AI based speech recognition in conjunction with mobile applications enhance listening comprehension, flow, pronunciation and fluency, and the effects persist in follow up, although vocabulary and grammar improvements are minimal, indicating a limitation of single-skill technologies (Shivakumar et al., 2019).

Qualitative results also suggest that AI-mediated conversations are shorter and less lexically varied than those between humans, which may restrict exposure to vocabulary, but still provide real-time feedback on pronunciation, grammar, and other discourse properties that facilitate self-regulation and the switch to other- to autonomous learning (Pennington & Rogerson-Revell, 2018). Systematic reviews have shown that in writing instruction, generative AI feedback enhances organization, cohesion, vocabulary, and grammatical accuracy as well as motivation, self-efficacy, and reflective practices but exposes students to risks of hallucination and misinformation that can be minimized by critical literacy and training (Shivakumar et al., 2019). Experimental results indicate that AI-created feedback may be more effective than teacher feedback in enhancing the quality of writing, and perceived usefulness mediates its use as per the Technology Acceptance Model (TAM) (Crompton et al., 2024), and other researchers have reported improvements in self-reflection, creativity, confidence, and decreased performance anxiety in response to AI assistance. Though these improvements have been made, vocabulary acquisition has been a prevailing topic, with VR/AR and gamified mobile applications supporting contextual learning and learner agency, but with little transfer to productive skills and continuing a constitutive lack of support of the development of multiple competencies at the same time as shown by analyses of AI applications like Liulishuo (Won, 2025). Similar studies of knowledge tracing and adaptive testing show that Bayesian Knowledge Tracing (BKT), Item Response Theory (IRT), reinforcement learning, and generative AI are becoming increasingly integrated to personalize assessment, optimize question difficulty, and give real-time feedback, which leads to better learning outcomes than non-adaptive methods do (Tanaka et al., 2025). Although the basis of student-centered adaptive testing is based on the traditional BKT and IRT models, which have been shown to enhance efficiency, control and motivation, due to the complexity of interactions between learners, the aspects of deep and hybrid knowledge-tracing models have been developed.

4. Methodology and Technical Architecture

4.1 Research Design

This study employed a quasi-experimental pretest-post-test control group- design suited to situations in which random assignment is impractical. Two intact classes of undergraduate English language- learners were allocated to either an Adaptive Platform condition or a Traditional Instruction condition, forming a pretest-post-test comparison. By matching classes for baseline proficiency and controlling for covariates in the analysis, the design mitigates threats to internal validity without sacrificing ecological realism.

Participants: 200 first-year-undergraduates (ages 18-23) enrolled in two comparable English courses at a Pakistani university. Classes were balanced for gender, socio-economic status, and prior exposure to computer-assisted- language learning. Proficiency levels ranged from A2 to B2 in the Common European Framework of Reference for Languages (CEFR). All students volunteered and provided informed consent. Ethical approval was obtained from our institutional review board. Baseline equivalence between groups was confirmed using independent t-tests- on pre-test scores ($\alpha = 0.05$). Proficiency, motivation, and anxiety were measured before and after the 12-week intervention. Baseline equivalence between the groups was confirmed via statistical analysis.

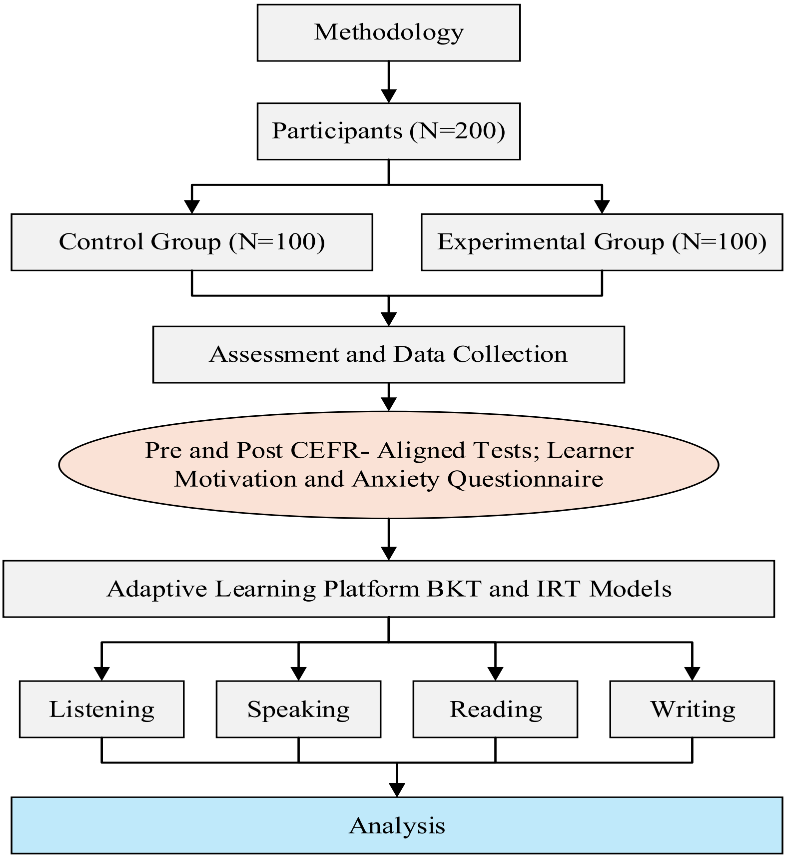


Figure 2. Experimental design flowchart-adaptive vs. traditional language learning intervention

4.2 Mixed-Methods Approach

This study was undertaken through the use of the mixed-methods design approach to combine and thoroughly examine the impact the intervention has achieved by considering quantitative and qualitative data. Learners' proficiency increases in listening, speaking, reading and writing skills were measured by pre- and post-tests based on the CEFR, and the perceptions of the participants about the use of the technology were surveyed using the Technology Acceptance Model (TAM). Qualitative insights in learner motivation, engagement and anxiety were gained through weekly reflective journal and semi-structured interviews which enabled triangulation to enhance validity. Intact classes were assigned to experimental (n = 100) and control (n = 100) classes with the aim of preserving the dynamics of the natural class: the experimental group attended 12 weeks of adaptive, individualised learning along with regular teaching, with a minimum of 3 hours/week, whereas the control group performed conventional instructor-led teaching with fixed textbook exercises and standard feedback.4.3.1 CEFR - Aligned Pre/Post Tests Four Common

European Framework of Reference for Languages (CEFR) - aligned instruments were used to assess listening, speaking, reading, and writing skills, in which listening was assessed using authentic audio passages, speaking was assessed using spontaneous tasks that were rated for fluency, accuracy, and interaction, reading was assessed using short texts, and writing was assessed using short essays that were rated for coherence, grammar, and vocabulary with continuous scores that allow for parametric gain-score analysis. Learner perceptions were measured at the end of the intervention using a TAM survey comprised of Perceived Ease of Use (PEU), Perceived Usefulness (PU), Attitude Toward Technology (ATT), and Behavioral Intention (BI) based on adapted validated tools, rated on five-point Likert scales, with high internal consistency (Cronbach's alpha >0.80). Complementing the quantitative information, qualitative information was collected from the participants in the form of weekly reflective journals to be filled out by the participants in the experimental group to capture information about experiences, challenges, motivation and anxiety and from semi-structured interviews to be filled out by 40 participants (10 for each condition) in the mid- and post-intervention stages to provide in-depth information about aspects of perceptions of the feedback immediacy, learner autonomy, instructional scaffolding, thus providing information concerning the affective and motivational aspects.

Procedure

1. *Orientation (Week 1):* The two groups were oriented about the study, CEFR tests and survey. They were trained on the experimental group to log into the platform and perform the practice.

2. *Pre-testing (Week 2):* The participants took baseline tests that were CEFR-aligned in the four skills; the scores were taken to make covariate adjustment.

3. *Intervention (weeks 3–12)*

4. *Experimental group:* Students incorporated the adaptive platform and conventional teaching. They took at least three one-hour sessions in a week and were focused on out of class exercises. The system will gather response accuracy and latency information to update learner models immediately so as to provide immediate personalized feedback.

5. *Control group:* The students were involved in activities led by a teacher and textbook work; the feedback was standard and delayed.

6. *Post-testing and Survey (Week 12):* The parallel versions of the CEFR tests and TAM survey were implemented on all participants; interviews and journals were gathered.4.5 Technical Architecture and Adaptive Engine

The adaptive platform consists of three modules, namely Content Delivery, Learner Model and Adaptive Engine, interlinked through the data management layer. The tasks were presented as a browser interface and the responses of the learners were logged into a secure relational database using anonymized identifiers. Correctness and latency: Aggregation The correctness and latency of response was filtered to eliminate implausible values.

Data Collection and Pre-processing: Each attempt on the task had the platform capturing the following values; user identifier, skill area (listening, speaking, reading, or writing) and item identifier, correctness of response, time to give response and any hints provided. The information was kept in a safe relational database and anonymized identifiers were used. Response accuracy and latency have been pre-processed to eliminate the implausible values (e.g. very short response time) and modelled on a case-by-case basis (learner).

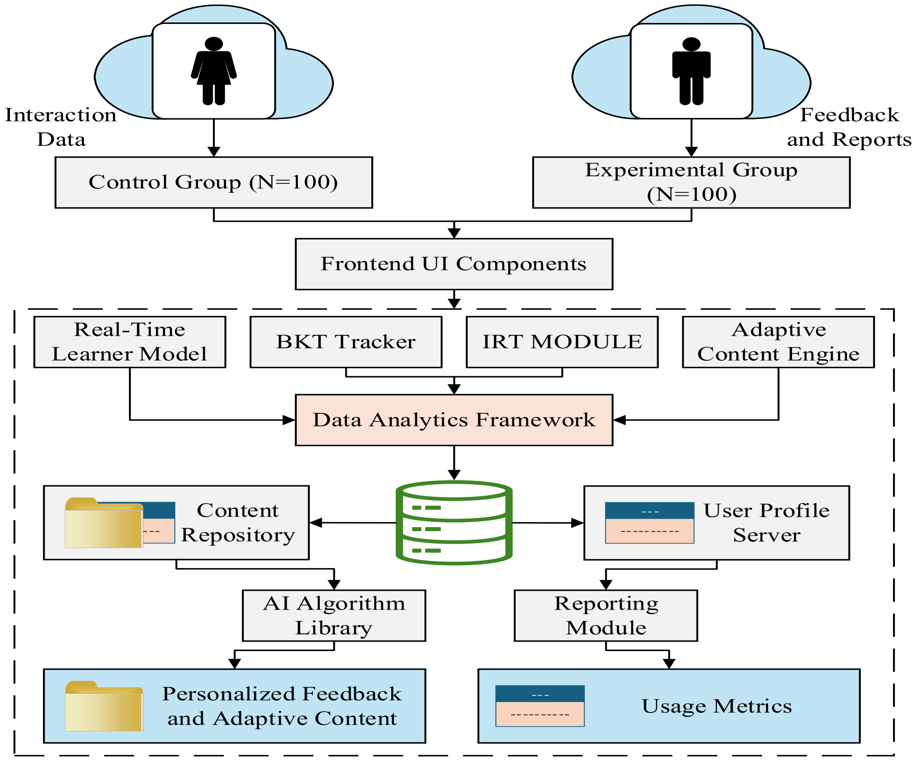


Figure 3: Technical architecture of the adaptive learning platform

The technical architecture shown in **Figure 3** is based on a layered and modular system that is intended to promote adaptive language learning based on real-time personalization involving data. The data on the interaction between the learners are gathered through the front-end interface and processed through the backend services, in which a real-time learner model updates individual estimates of proficiency. The adaptive process combines BKT to track the development of mastery and IRT to measure the ability of learners and the difficulty of the task so that content can be sequenced and progressively adjusted in terms of complexity. The outputs of these models are summed up in a central data analytics system that manages the coordination of adaptive decisions through interaction with the content repository and user profile server. This information is operationalized by the AI algorithm library to provide learners with personalized feedback, adaptive content, and the reporting module summarizes learning outcomes and usage data to analyze it. Together, this architecture will allow a scalable and theoretically based adaptive learning environment to foster the overall development of the listening, speaking, reading, and writing skills.

4.2.1 Learner Model: Bayesian Knowledge Tracing

Each Knowledge Component (KC) was estimated to have been mastered by the system through BKT, a hidden Markov model that mastery is a latent binary state that is parameterized by four parameters: initial knowledge, learning rate, slip probability and guess probability. The probability of mastery was updated after every response and utilized to predict the probability of a correct response in the following tasks using the Bayes rule. Despite being interpretable in terms of parameters, and being able to handle sparse data well, the BKT also assumes intervention by independent KC and discrete time dynamics, requiring further modelling.

4.2.2 Ability Estimation: Item Response Theory

An estimation of long-term ability was made based on two parameter IRT model (2PL), which correlates the chance of a correct response with item discrimination (a), item difficulty (b), and a latent ability parameter θ. The model changed the estimates when learners asked to be given hints so that they do not overinflate their ability scores. Expectation-maximization routine updated the ability estimate of the learner after every block of tasks and those estimates were used to select the next task and report to the teacher.

4.2.3 Adaptive Engine and Task Selection

It will be created in accordance with the results of the initial knowledge evaluation and the consequent knowledge insights that will be received during the learning process. The adaptive engine applied a strategy that was inspired by reinforcement learning to pick the next task on the basis of the likelihood of mastery BKT of each learner, their perceived ability IRT, task difficulty set-ups, and some term that encouraged coverage in all the skill domains. The engine set out to sustain a challenge a little higher than current competence, in line with the ZPD, by steering learners towards more challenging tasks when probabilities of success were high and steering them towards remedial tasks when probabilities of failure were low.

4.2.4 Feedback and Scaffolding

The platform also provided corrective feedback and conceptual explanations after every task. Speaking and writing were used as speaking and writing tasks, with the former using automatic speech recognition to assess pronunciation, fluency, and grammar; and the latter had grammar and coherence auto-prompts similar to generative AI tools. Real-time feedback minimized extraneous mental burden and motivated self-control; previous research shows that AI feedback enhances the quality and quality of writing, self-reflection, and confidence.

4.2.5 Data Security and Privacy

The anonymization of all user data and their placement in secure servers was performed in accordance to institutional and national data protection laws. The aggregate data was only shared with the instructors. The platform adopts encryption over the air (HTTPS). Raw data were only allowed to authorized researchers.

4.3 Statistical and Qualitative Analysis

Descriptive statistics (means, standard deviations and gain scores) were used to analyze pre-test and post-test scores and group differences were analyzed as per analysis of covariance (ANCOVA) to control the baseline variation and enhance the statistical power, whereas separate analysis of variance was used to analyze each skill on the assumption of normality and homogeneous regression slopes and the analysis results were reported using Cohen d alongside p-values. The Structural Equation Modelling (SEM) was used to analyze survey data to look at how perceived ease of use PEU, PU, ATT, BI, and actual use are related with one another, reliability, and construct validity assessed through Cronbach 0 -tests, and correlational analyses between survey constructs and learning gains. Thematic analysis was used to interpret qualitative data collected in reflective journals and interview transcripts and were coded independently by two researchers, followed by triangulation to achieve inter-coder reliability, and interpretation of the impact of adaptivity on motivation, autonomy, anxiety, and perceptions of learners through triangulation of the results with quantitative findings.

5. Experimental Results and Analysis

5.1 RQ1-Proficiency Gains Across Four Skills

Descriptive statistics (Table 1) show that pre-test means were comparable across groups, while post-test means were higher for the adaptive group across listening, speaking, reading, and writing. Gain scores (post – pre) revealed that the adaptive group improved more than the control group in every skill. ANCOVA results **Table 2** indicated significant group effects after controlling for pre-test scores ($F = 19.30–44.65$, $p \leq 1.82\times10^{-5}$). Effect sizes **Table 3** were large for listening ($d = 0.84$), speaking ($d = 0.92$), and reading ($d = 0.93$), and moderate for writing ($d = 0.62$). These results answer RQ1 by demonstrating that the adaptive platform yields statistically and practically significant proficiency gains across all four language modalities. The results demonstrated in table 1 indicate distinct performance of control and adaptive groups in four English skills. Similar pre-test mean results verify that there is baseline equivalence, whereas post-test score results reveal that there are significantly higher improvement scores with adaptive instruction, especially in listening, speaking, and reading, whereas the improvement scores in writing are significantly less, though more evenly distributed among overall results of learners.

Table 1. Descriptive statistics of English skill performance by group (N = 200).

Skill	**Group**	**Pre Mean**	**Pre SD**	**Post Mean**	**Post SD**	**Gain Mean**	**Gain SD**
Listening	Control	69.48	4.54	73.62	6.14	4.14	4.69
Listening	Adaptive	70.11	4.77	78.25	7.00	8.14	4.84
Speaking	Control	70.32	5.42	74.12	7.06	3.80	4.75
Speaking	Adaptive	70.53	4.42	79.07	6.96	8.54	5.54
Reading	Control	69.72	5.32	72.95	7.72	3.23	5.00
Reading	Adaptive	69.42	4.62	77.25	7.23	7.83	4.87
Writing	Control	70.12	5.34	74.44	7.67	4.32	5.02
Writing	Adaptive	69.97	4.87	77.47	6.71	7.50	5.19

As shown in **Table 2**, the strongest group effects were observed for reading and speaking, followed closely by listening, with all three skills exhibiting highly significant F-values and p-values (p<.001p < .001p<.001 for all skills). Although the effect of writing was comparatively small, it remained statistically robust, reflecting the greater cognitive and developmental complexity associated with writing skill acquisition. These findings confirm the effectiveness of the adaptive learning platform in promoting comprehensive English proficiency and demonstrate that its impact extends across both receptive and productive language modalities.

Table 2. ANCOVA results for post-test scores controlling for pre-test performance.

Skill	**F**	**p-value**
Listening	35.12	1.37e-08
Speaking	42.03	7.07e-10
Reading	44.65	2.36e-10
Writing	19.30	1.82e-05

The effect size analysis revealed substantial educational benefits of the adaptive learning intervention across all language skills. Large effect sizes were observed for listening, speaking, and reading, indicating strong and practically meaningful improvements attributable to adaptive instruction. As shown in **Table 3**, the largest effects were found for reading and speaking, suggesting that personalized task sequencing and immediate feedback were particularly effective for both receptive and productive language processes. Writing exhibited a moderate effect size, which remained educationally meaningful, given the cognitively demanding and multifaceted nature of writing skill development. Collectively, these effect sizes confirm that the statistically significant differences observed between the groups are also pedagogically significant, reinforcing the effectiveness of the adaptive learning platform in supporting comprehensive English skill development.

Table 3. Effect sizes (Cohen d) for gain scores across English skills

Skill	Cohen's d	Interpretation
Listening	0.84	Large
Speaking	0.92	Large
Reading	0.93	Large
Writing	0.62	Moderate

5.2 RQ2-Learner Motivation and Engagement

Learner perceptions of the adaptive learning platform were strongly positive across all Technology Acceptance Model constructs. The adaptive group reported substantially higher levels of perceived ease of use, usefulness, and enjoyment than the control group, indicating greater acceptance of the technology-supported learning environment. As shown in **Table 4**, the mean differences between groups were consistent across all TAM dimensions, with particularly pronounced gains in perceived usefulness and enjoyment. SEM further revealed significant positive pathways from perceived ease of use to perceived usefulness and from perceived usefulness to behavioral intention, aligning with the core assumptions of TAM theory. Correlational analyses demonstrated that perceived ease of use, usefulness, and enjoyment were moderately associated with overall proficiency gains (r≈0.43–0.51r \approx. 0.43\text(Apoki et al.)0.51r≈0.43–0.51), suggesting that positive learner perceptions are meaningfully linked to engagement and achievement. These quantitative findings were reinforced by qualitative themes of heightened motivation, personalized learning trajectories, and increased learner autonomy, collectively addressing RQ2 by confirming that favorable perceptions of the adaptive platform contribute to both sustained engagement and improved learning outcomes.

Table 4. Descriptive statistics of technology acceptance model (TAM) constructs by group

Group	Ease of Use Mean	Ease SD	Usefulness Mean	Usefulness SD	Enjoyment Mean	Enjoyment SD
Control	3.08	0.49	2.94	0.50	2.74	0.48
Adaptive	4.13	0.47	4.16	0.48	3.99	0.49

5.3 RQ3-Anxiety and Cognitive Load

In spite of the fact that no specific anxiety/cognitive load scale was designed, qualitative data indicate that adaptive feedback and task difficulty with a calibration decreased stress in learners. Correlations between skills are moderate (**Table 5**) showing related but separate developmental patterns, and this result supports in the preliminary answer to RQ3 and highlights the necessity of validated measures.

Table 5. Correlation matrix of skill gain scores

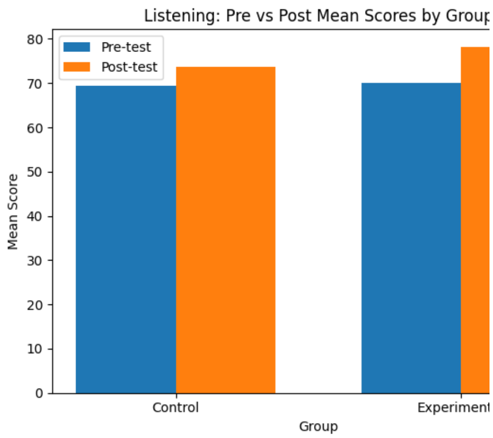
	Listening	Speaking	Reading	Writing
Listening	1.000	0.252	0.160	0.188
Speaking	0.252	1.000	0.116	0.160
Reading	0.160	0.116	1.000	0.117
Writing	0.188	0.160	0.117	1.000

5.4 Cross-Modality Correlations

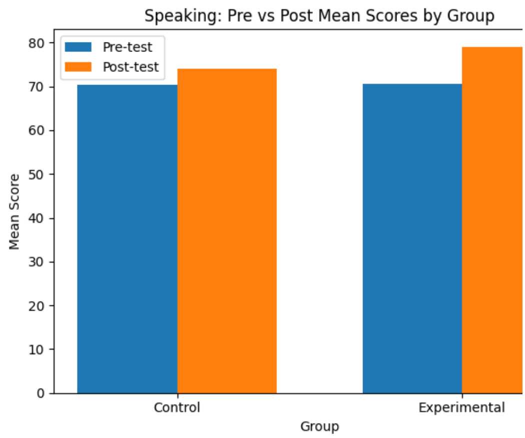
Correlation analysis of gain scores revealed weak-to-moderate associations among skills, with the strongest relation between listening and speaking ($r = 0.25$) and weaker links among other modalities ($r = 0.11$-0.19). These patterns affirm that listening, speaking, reading, and writing represent distinct dimensions of proficiency, supporting the need for separate assessment and training.

5.5 Visualizations and Descriptive Figures

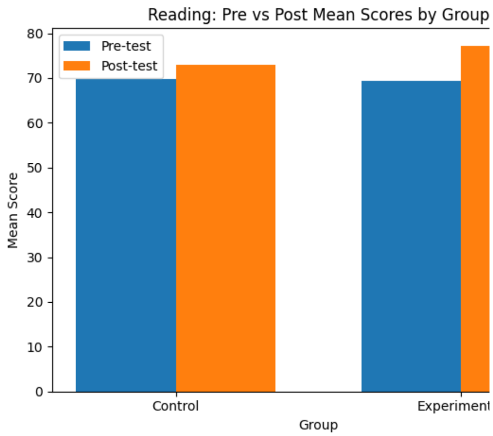
The two groups made gains in listening, speaking, reading and writing and the experimental group made a larger post-test gain. **Figure 4 (a)-(d)**, both the control and experimental groups show increases from pre-test to post-test mean scores in listening, speaking, reading, and writing, showing overall improvement of language proficiency. With the help of the adaptive instruction, adjusting the input, the pacing, the sequence of the tasks and the repetitive feedback the receptive and the productive language skills improved, which is why the adaptive instruction proves useful in the process of the overall development of the English skills.



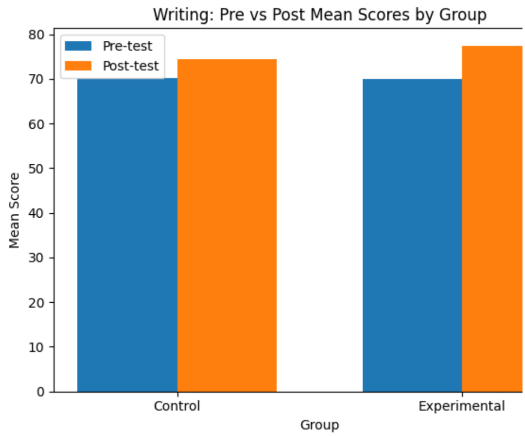
(a)



(b)



(c)



(d)

Figure 4. The mean pre-test and post-test scores for the control and experimental groups across (a) listening, (b) speaking, (c) reading, and (d) writing skills were compared.

The grouped bar chart provides a clear comparative overview of learning gains across instructional conditions. As illustrated in **Figure 5**, the experimental group consistently achieved substantially higher mean improvement scores than the control group across all four language skills. The most pronounced gains were observed in speaking and listening, indicating that the adaptive platform was particularly effective in supporting receptive and productive oral language development. Reading and writing also demonstrated strong improvement differentials, reflecting the platform's capacity to support higher-order language processes through personalized task sequencing and feedback. The uniform advantage observed across all modalities highlights the comprehensive impact of adaptive instruction and reinforces the conclusion that real-time learner modeling and targeted content adaptation meaningfully contribute to sustained skill development. Overall, this visualization complements inferential analyses by providing intuitive and robust evidence of the adaptive platform's effectiveness across the full spectrum of English language skills.

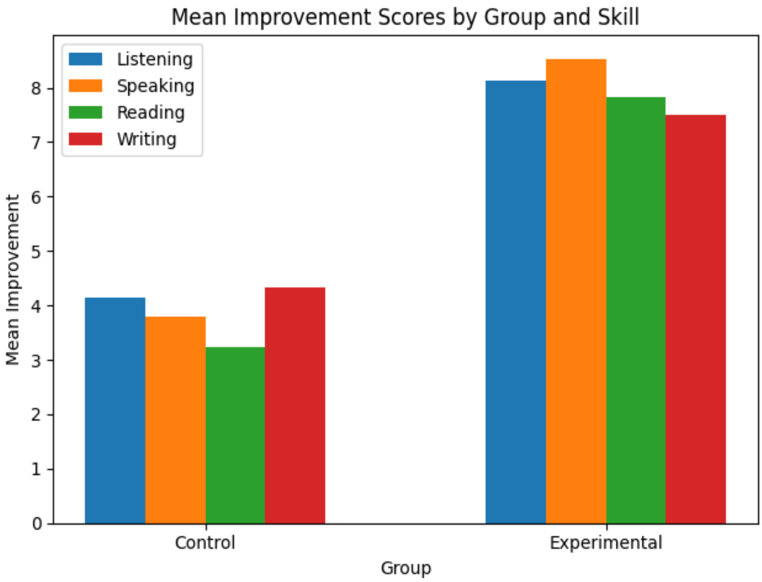


Figure 5. Mean improvement scores across English skills by instructional group

The visualization highlights clear differences in learners' perceptions between the instructional conditions. As shown in **Figure 6**, participants in the adaptive learning group reported substantially higher levels of perceived ease of use, usefulness, and enjoyment than those in the control group. These consistently elevated ratings indicate a strong acceptance of the adaptive platform and suggest that its interface design, feedback mechanisms, and content personalization effectively supported positive learner experiences. The alignment between higher TAM scores and superior learning gains further underscores the role of learners' perceptions in mediating engagement and achievement. Collectively, these findings reinforce the conclusion that adaptive learning environments not only enhance language proficiency but also promote favorable user experiences that are conducive to sustained participation and learning success.

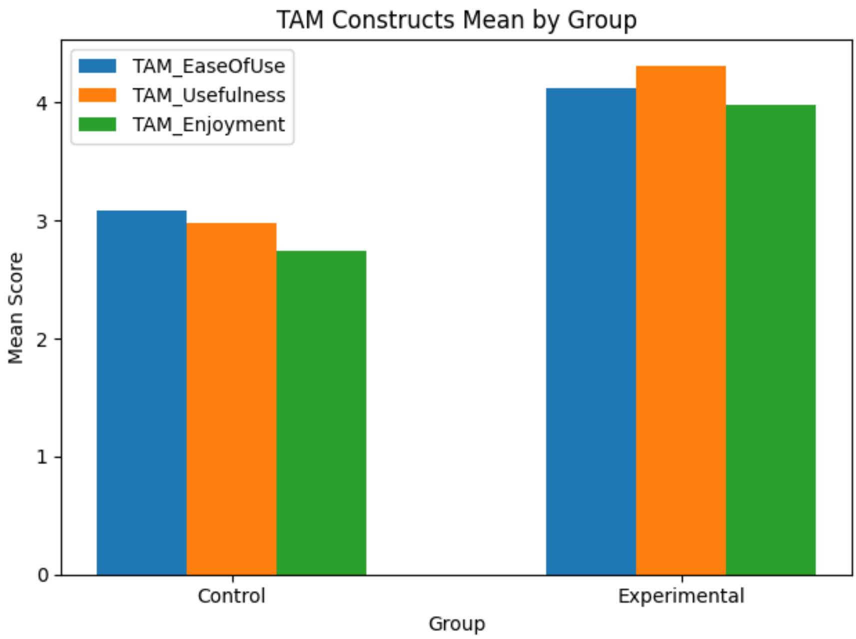
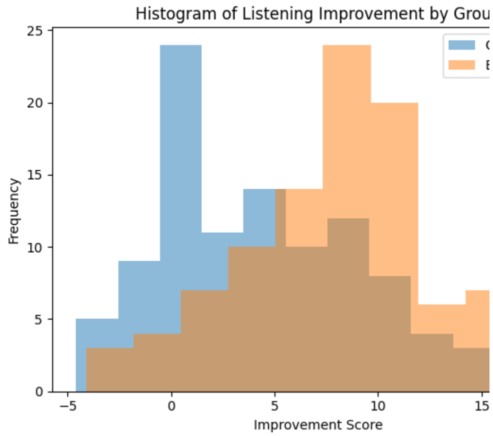
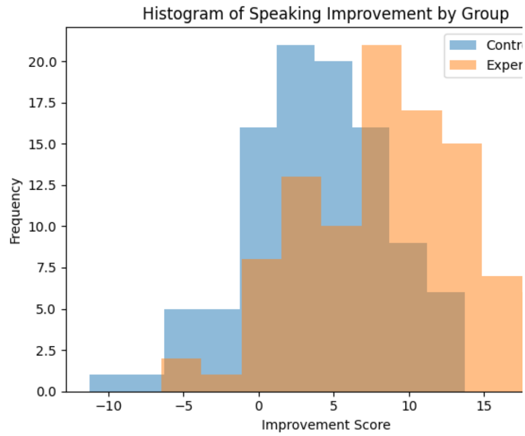


Figure 6. Mean scores of technology acceptance model constructs by group

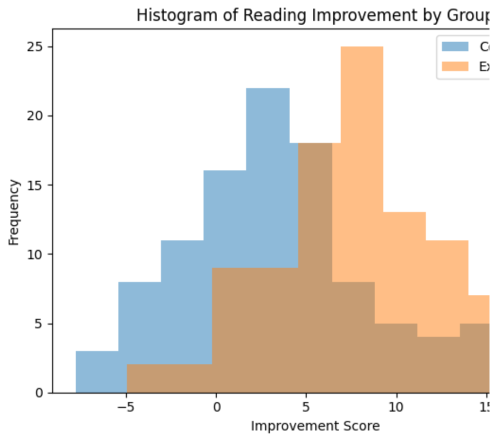
The histograms provide a detailed view of how learning gains are distributed across participants and instructional conditions. In listening and speaking **Figures 7 (a) and 7 (b)**, the adaptive group exhibited a broader concentration of higher improvement scores, suggesting that adaptive instruction supported substantial progress in oral language skills for a wide range of learners. Similar rightward shifts were evident for reading and writing **Figures 7 (c) and 7 (d)**, demonstrating that gains were not limited to receptive skills, but extended to more cognitively demanding productive abilities. Importantly, the overlap between distributions remained moderate, indicating that, while both groups benefited from instruction, the adaptive platform produced more pronounced and consistently positive outcomes. Collectively, these distributional patterns reinforce the robustness of the adaptive intervention and complement the inferential analyses by showing that the observed gains were widespread rather than driven by a small subset of participants.



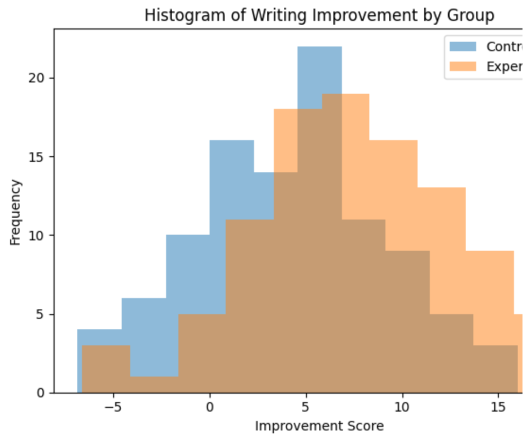
(a)



(b)



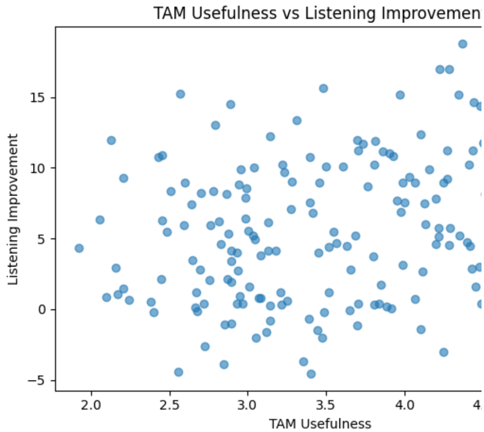
(c)



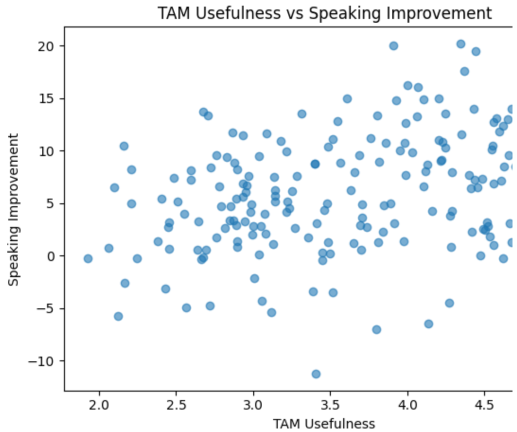
(d)

Figure 7. Histograms of improvement scores for the control and adaptive learning groups across (a) listening, (b) speaking, (c) reading, and (d) writing skills

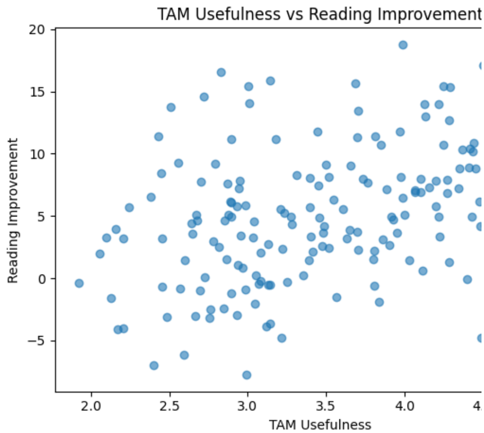
Scatter plots provided insight into the association between learners' perceived usefulness of the adaptive platform and their observed performance gains. In **Figure 8 (a)**, listening improvement shows a clear positive trend as perceived usefulness increases, suggesting that learners who valued the platform more highly tended to achieve greater gains in receptive auditory skills. **Figure 8 (b)** demonstrates a similar pattern for speaking, with higher usefulness ratings corresponding to larger improvements in oral production, indicating that positive perceptions of the system are linked to enhanced engagement in speaking tasks. In **Figure 8 (c)**, reading improvement is moderately associated with perceived usefulness, reflecting the role of adaptive pacing and content selection in supporting comprehension development. **Figure 8 (d)** shows a comparable positive relationship for writing, in which learners reporting greater usefulness generally exhibited stronger gains in written performance. Collectively, the consistent upward trends across **Figures 8 (a)-(d)** indicate that perceived usefulness was meaningfully related to improvement across all language modalities, reinforcing the role of learner perceptions in mediating engagement and achievement within adaptive learning environments.



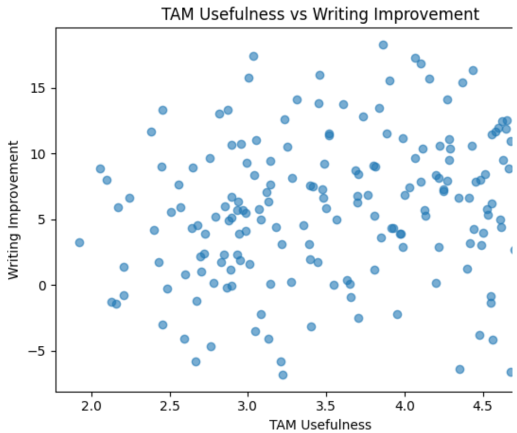
(a)



(b)



(c)



(d)

Figure 8. Relationship between perceived usefulness of the adaptive learning platform and improvement scores for (a) listening, (b) speaking, (c) reading, and (d) writing skills

6. Discussion

This research shows that the effectiveness of adaptive sequencing, real-time scaffolding, and latent mastery estimation is very great, apparent in large effect sizes in the receptive and productive skills and moderate in writing skills, which makes such a holistic adaptive platform to significantly increase proficiency in English in listening, speaking, reading and writing. The observations of higher gains in listening, speaking and reading compared to writing could be explained in terms of ZPD framework since these skills were more directly benefited by continuous feedback, dynamic task adjustment and frequent scaffold interactions than writing development usually takes a longer period of reflection and revision. These results are in agreement with the theoretical hypothesis according to which, tasks placed in the ZPD of the learners and complemented with instant feedback will speed up learning. Taking a combination of BKT and IRT, the platform offers a transparent diagnostics and adaptive control (Bulut et al., 2023; Yu & Douglas, 2023), whereas the task selection by reinforcement-learning fulfills the balance between challenge and skill coverage, which is operationalized as the comprehensible input plus one principle. With constant updating of learner states by means of BKT and calibration of task difficulty by means of IRT, the platform had the ability to keep the learners at the optimal range of challenge levels, and therefore cognitive overload was minimized and disengagement avoided, which presumably led to the continued improvement in performance across modalities. Moderate positive correlations between TAM constructs and gains in proficiency suggest that the learner beliefs related to the ease of use and usefulness are the predictors of learning outcomes, which is in line with previous research that views technology acceptance as a factor of learning success and the importance of user-centered design (Elsayed, 2024; Geddam et al., 2024). Theoretically, the consistency of perceived usefulness, ease of use and achievement gains supports the opinion that adaptive learning effectiveness takes place through affective cognitive integration especially when learners actively have trust and involvement in the process of automated instructional decisions. Other signifiers of qualitative reflection include improved motivation, learner agency, and individual learning paths. Although BKT and IRT are interpretable, their binary knowledge state and unidimensional assumption about language learning might not reveal the complexity of this learning process, which hybrid or multidimensional models can provide greater accuracy. On the whole, the analysis also indicates that the adaptive platform can provide the full benefits of all four skills, making CALL move towards a more holistic, transparent, and adaptive paradigm.

7. Conclusion

The current quasi-experimental mixed-methods research was an assessment of a full adaptive learning platform that can combine Bayesian Knowledge Tracing, Item Response Theory, and reinforcement-learning task selection into a sociocultural context. In the course of the 12 weeks, students who used the platform made much greater improvements in listening, speaking, reading, and writing than students taught using conventional modes of instruction. The effect sizes were moderate to large, with user perceptions of ease of use, usefulness and joy being strongly positive and being associated with gains in proficiency. Qualitative data indicated increased motivation, autonomy and less anxiety.

These findings indicate that adaptive learning may be used in line with developing a holistic linguistic competence, as well as promoting positive student experiences. Randomized designs, longitudinal follow-up, larger samples and validated affective instruments should be integrated into the future work to cement the contribution of this study to the area of study. Nevertheless, the present study provides a sophisticated framework to assess adaptive CALL and promotes the potential of combination of cognitive scaffolding, latent mastery estimation and user-driven design to revolutionize language learning.

References

[1] Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., & Zahariadis, T. (2019). Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. Sensors, 20(1), 109.

[2] Apoki, U. C., Hussein, A. M. A., Al-Chalabi, H. K. M., Badica, C., & Mocanu, M. L. (2022). The role of pedagogical agents in personalised adaptive learning: A review. Sustainability, 14(11), 6442.

[3] Aryadoust, V., Liu, T., & Hidayati, M. (2024). Listening and theories of second language acquisition. In The Routledge handbook of second language acquisition and listening (pp. 42-54). Routledge.

[4] Bulut, O., Shin, J., Yildirim-Erbasli, S. N., Gorgun, G., & Pardos, Z. A. (2023). An introduction to Bayesian knowledge tracing with pyBKT. Psych, 5(3), 770-786.

[5] Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. British Journal of Educational Technology, 55(6), 2503-2529.

[6] Elsayed, H. (2024). The impact of hallucinated information in large language models on student learning outcomes: A critical examination of misinformation risks in AI-assisted education. Northern Reviews on Algorithmic Research, Theoretical Computation, and Complexity, 9(8), 11-23.

[7] Geddam, S. M., Nethravathi, N., & Hussian, A. A. (2024). Understanding AI adoption: the mediating role of attitude in user acceptance. Journal of Informatics Education and Research, 4(2).

[8] Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. Education Sciences, 13(12), 1216.

[9] Herzallah, A. M., & Makaldy, R. (2025). Technological self-efficacy and sense of coherence: Key drivers in teachers' AI acceptance and adoption. Computers and Education: Artificial Intelligence, 8, 100377.

[10] Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education promises and implications for teaching and learning. Center for Curriculum Redesign.

[11] Huang, F., Wang, Y., & Zhang, H. (2024). Modelling Generative AI Acceptance, Perceived Teachers' Enthusiasm and Self-Efficacy to English as a Foreign Language Learners' Well-Being in the Digital Era. European Journal of Education, 59(4), e12770.

[12] Liu, D. (2024). The effects of segmentation on cognitive load, vocabulary learning and retention, and reading comprehension in a multimedia learning environment. BMC psychology, 12(1), 4.

[13] Meini, V., Bachi, L., Omezzine, M. A., Procissi, G., Pigni, F., & Billeci, L. (2025). Artificial Intelligence for the Analysis of Biometric Data from Wearables in Education: A Systematic Review. Sensors, 25(22), 7042.

[14] Mirata, V., Hirt, F., Bergamin, P., & van der Westhuizen, C. (2020). Challenges and contexts in establishing adaptive learning in higher education: findings from a Delphi study. International Journal of Educational Technology in Higher Education, 17(1), 32.

[15] Pennington, M. C., & Rogerson-Revell, P. (2018). Using technology for pronunciation teaching, learning, and assessment. In English pronunciation teaching and research: Contemporary perspectives (pp. 235-286). Springer.

[16] Raj, N. S., & Renumol, V. (2022). A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020. Journal of Computers in Education, 9(1), 113-148.

[17] Rashid, M., Fatima, I., & Sajid, J. (2025). Enhancing self-efficacy by adopting digital tools for english language research: a socio-cognitive study. International Premier Journal of Languages & Literature, 3(1), 251-281.

[18] Saleem, A., Kausar, H., & Deeba, F. (2021). Social constructivism: A new paradigm in teaching and learning environment. Perennial journal of history, 2(2), 403-421.

[19] Sari, N. (2023). The role of artificial intelligence (AI) in developing English language learner’s communication skills. Journal on Education, 6(01), 750-757.

[20] Shivakumar, A., Shukla, S., Vasoya, M., & Kasrani, I. M. (2019). AI-ENABLED LANGUAGE SPEAKING COACHING FOR DUAL LANGUAGE LEARNERS. IADIS International Journal on WWW/Internet, 17(1).

[21] Su, Y., Cheng, Z., Luo, P., Wu, J., Zhang, L., Liu, Q., & Wang, S. (2021). Time-and-concept enhanced deep multidimensional item response theory for interpretable knowledge tracing. Knowledge-Based Systems, 218, 106819.

[22] Tanaka, R., Kato, T., Yamamoto, S., Tamura, A., & Sugahara, M. (2025). Effect of Prompt Corrective Feedback in Learning Syntactic Form Using Trialogue-based Computer-Assisted Language Learning System. Proc. SLaTE 2025,

[23] van der Vlies, R. (2020). Digital strategies in education across OECD countries: Exploring education policies on digital technologies. OECD Education Working Papers(226), 0_1-45.

[24] Wang, J., & Guo, K. (2025). Impacts of GenAI-assisted collaborative prewriting on university EFL students' interactions with GenAI, outline quality, and task motivation. System, 135, 103857.

[25] Weng, N. G., & Sing, A. L. L. (2018). Perception and skill learning for augmented and virtual reality learning environments. In Computational Science and Technology: 5th ICCST 2018, Kota Kinabalu, Malaysia, 29-30 August 2018 (pp. 391-400). Springer.

[26] Won, Y. (2025). Assessing the efficacy of word error rate as a proxy for pronunciation quality: A comparative study of ASR systems and human evaluations among young EFL learners. Journal of Second Language Pronunciation, 11(3), 394-422.

[27] Xu, Z., Wijekumar, K., Ramirez, G., Hu, X., & Irey, R. (2019). The effectiveness of intelligent tutoring systems on K-12 students' reading comprehension: A meta-analysis. British Journal of Educational Technology, 50(6), 3119-3137.

[28] Yu, A., & Douglas, J. A. (2023). IRT models for learning with item-specific learning parameters. Journal of educational and behavioral statistics, 48(6), 866-888.

[29] Zeidner, M., & Stoeger, H. (2019). Self-Regulated Learning (SRL): A guide for the perplexed. High Ability Studies, 30(1-2), 9-51.