10.48047/jocaaa.2024.33.08.249

Personalized Language Learning with AI: An Intelligent Companion Approach

Anil Khatak[1]
ORCID ID-0000-0001-9103-0210
Department of Allied Health Sciences, Guru Jambheshwar University of Science and Technology, Hisar, India, 125001
khatak10anil@gmail.com

Vikas Kajla[2]
ORCID ID-0009-0004-0218-877X
Department of Allied Health Sciences, Guru Jambheshwar University of Science and Technology, Hisar, India, 125001
kajla.vikas27@gmail.com

Sheenam Naaz[3]
Department of Computer Science and Engineering, SSCSE, Sharda University,
Greater Noida, India, 201310
naazsheenam.work@gmail.com

Abstract

The intelligent companion is an advanced AI-powered language learning companion designed to transform the way users acquire language skills through immersive, interactive, and personalized virtual experiences. The project bridges the gap between traditional language education and modern AI capabilities by integrating Large Language Models (LLMs) with 3D environments and avatars to create an engaging, adaptive, and real-time learning experience. The core philosophy of the work is that language learning doesn't have to be boring. This is not necessarily a matter of passively absorbing information. The best way to learn a language is through exposure to new words in context, through the dynamic medium of conversation, and with the help of the visual social cues that help to understand every language, including intonation, gestures, and expressions. Dedicated to transforming the abstract concepts of language into lived experiences, the aim of this research is to provide users the chance to immerse themselves in the language as part of their daily lives, rather than just learning it.

Keywords: Real-Time learning, Gesture, Artificial Intelligence (AI), Large Language Models (LLMs)

1. Introduction

Learning a new language is more than memorizing vocabulary or mastering grammar rules. It involves developing the confidence to communicate in real-life situations with natural fluency. Many existing digital language learning tools are accessible but still rely on scripted dialogues, repetitive drills, and static lesson structures, which can limit learner engagement and contextual understanding. Recent developments in artificial intelligence (AI) and large language models (LLMs) have created opportunities to transform how learners acquire, practice, and retain language skills [3], [7].

10.48047/jocaaa.2024.33.08.249

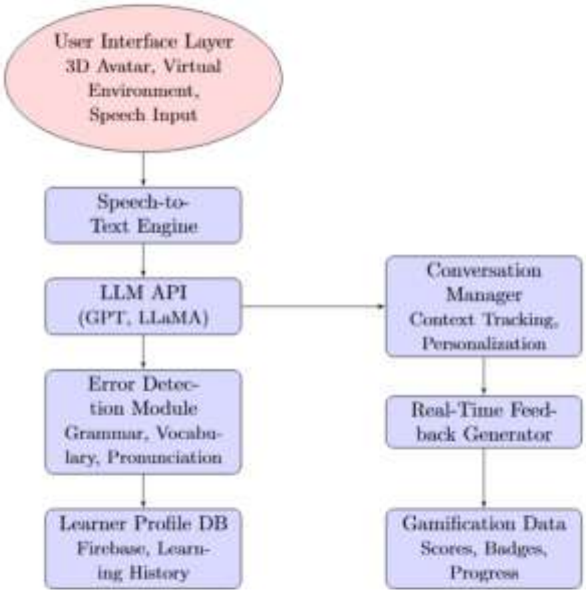

Fig. 1. Proposed architecture of the AI-powered personalised language learning companion

2. Contribution and Novelty

Human–computer interaction research shows that visual cues, gamification, and adaptive feedback can improve learner motivation and retention [4,6]. By embedding conversational practice into realistic virtual scenarios such as ordering food in a café or participating in a meeting the system replicates the social and cognitive aspects of honest communication. Studies further indicate that contextual learning paired with timely feedback can significantly improve long-term retention and learner confidence [5,14].

The key contributions of this work are as follows:

1. The system integrates lifelike avatars with gestures, facial expressions, and environment-aware interactions to create an immersive 3D learning environment.
2. It uses transformer-based LLMs to provide context-sensitive feedback and dynamically adjust the difficulty of interactions in real time.
3. It incorporates gamified learning pathways with rewards and scenario-based challenges to sustain learner motivation and track progress through performance analytics.

For reference, Figure 1 (System Architecture) shows the high-level vision of the proposed work. The system aims to address the gaps of earlier language learning platforms by combining transformer-based LLMs with expressive 3D avatars and gamified pathways. While previous systems focused primarily on text-based or voice-only feedback [10,16], this approach integrates speech recognition, contextual scene rendering, and personalized learning progression.

Learners interact with a virtual tutor that listens and responds while adapting to their proficiency level, emotional cues, and learning goals [11].

3. Literature Survey

To build an advanced, immersive language learning solution, a comprehensive review of existing technologies, educational frameworks, and AI-based applications was undertaken. The survey includes both commercial platforms and research-based solutions in EdTech, focusing on their strengths, limitations, and how our work innovates beyond them. Table 1 (Literature Survey) summarizes the work of several reviewed studies, highlighting their key contributions, limitations, and gaps.

10.48047/jocaaa.2024.33.08.249

Duolingo

Duolingo is one of the most widely adopted language learning applications, offering gamified lessons with increasing difficulty. It uses a reward system (XP, streaks, badges) to promote consistency. However, the learning experience is heavily text/image-based, with limited audio interaction and almost no focus on real-time conversation. Studies [22] have evaluated its effectiveness in developing both receptive and productive language skills, finding strengths in vocabulary acquisition but limitations in conversational fluency.

- Predefined responses limit creative expression.
- Lacks real-world conversational immersion.
- AI feedback is limited to right/wrong answers without personalized corrections.

MondlyVR

Mondly introduced Virtual Reality (VR) to language learning through interactive scenarios such as airport check-ins, restaurants, or office settings. The VR-based approach adds visual and spatial context to vocabulary acquisition. Recent work [23] highlights its potential for immersive learning but notes that its adaptiveness is still limited compared to AI-driven conversation models.

- No LLM-powered grammar/syntax correction.
- VR responses are scripted and not truly adaptive.
- Limited user-specific learning paths.

ELSA Speak

ELSA focuses on accent training and pronunciation using speech recognition and phonetic analysis. It is especially helpful for non-native English speakers preparing for IELTS/TOEFL. Research [24] has shown that ELSA can significantly improve learners' pronunciation accuracy, although it remains limited in broader language skills.

- No avatar-based visual cues.
- No grammar correction or context-based responses.
- Cannot support long-form conversation practice.

Pratiksha.ai

An emerging tool using natural language processing (NLP) to simulate job interviews or casual conversations. It generates semi-custom dialogues based on user input and provides basic correction. It is listed among the latest AI language learning tools in [25], but its current version still lacks advanced immersive features.

- Text-based feedback only.
- No immersive 3D environments.
- Minimal use of gestures or emotion-rich responses.

	Methodology / Approach	**Dataset / Context**	**Key Contribution**	**Performance Metrics**	**Limitations**
[1]	Adaptive hypermedia and educational systems modelling	Various adaptive web systems	Framework for adaptive personalization in learning	Not specified	General framework, not language-specific

10.48047/jocaaa.2024.33.08.249

[2]	Information retrieval models for personalization	Text corpora	Core IR principles for language content retrieval	Retrieval accuracy	Not language-learning specific
[3]	ASSISTments platform for adaptive learning	K–12 students	Scalable adaptive tutoring ecosystem	Student performance gain	Primarily math-focused
[4]	NLP for CALL systems	ESL/EFL learners	Integration of NLP in language learning tools	Learner engagement	Limited real-time adaptability
[5]	Bayesian Knowledge Tracing (BKT)	Simulated learner data	Procedural knowledge modelling	Prediction accuracy	Assumes binary mastery
[6]	Tutoring systems review	Multiple domains	Analysis of ITS behavior and design	Comparative metrics	Not AI-specific
[7]	Adaptive vocabulary learning via AI	L2 learners	AI-based vocabulary personalization	Vocabulary retention rate	Limited to vocabulary
[8]	Educational tech trend analysis	EdTech case studies	Horizon scanning for AI in learning	Trend mapping	No experimental validation
[9]	Dialog-based ITS for language	ESL learners	Conversational tutoring design	Learner satisfaction	Small-scale pilot
[10]	Chatbots for L2 speaking practice	EFL students	AI chatbot integration in speaking tasks	Speaking fluency gains	Limited natural conversation depth
[11]	Machine learning for individualization	Online student logs	Personalized ML recommendation	Prediction accuracy	Requires large datasets
[12]	Adaptive practice scheduling	L2 learners	Personalized scheduling for retention	Retention improvement	Narrow task scope
[13]	Adaptive reading recommendation	Language learners	Personalized reading material selection	Engagement rate	Limited genre diversity
[14]	User-centered design of AI companion	Prototyping with learners	Companion interface tailored to learners	Usability scores	Prototype stage only
[15]	Constraint-based tutoring	Multiple domains	Error-detection without domain-specific rules	Accuracy of error feedback	Limited adaptivity
[16]	Review of ASR in CALL	Speech recognition datasets	ASR evaluation in oral language practice	Recognition accuracy	Accent bias
[17]	Mobile-assisted language learning review	Mobile learning apps	Overview of MALL trends	Adoption metrics	Mostly descriptive
[18]	Adaptive error correction feedback	AI tutoring system logs	Real-time feedback personalization	Error correction success rate	Limited to grammar tasks

10.48047/jocaaa.2024.33.08.249

[19]	MALL special topic review	Mobile-based learning contexts	Identifies mobile learning opportunities	Case study synthesis	Lacks experimental data
[20]	Personalized AI companion design study	University learners	Engagement-focused AI companion	Learner engagement scores	No longitudinal testing
[21]	Conversational agents for vocabulary	Longitudinal L2 learning	Chatbot-based vocabulary training	Vocabulary gain rate	Limited context awareness

4. Implementation Methodology

The main problem with traditional language learning apps is that they offer limited personalization and lack realistic interaction. They struggle with engagement, contextual learning, and fluency development.

LinguoMate addresses these issues by merging 3D immersion, AI feedback, and gamified conversation practice.

When developing the intelligent companion, industry best practices and modern software engineering principles were followed. The development approach was divided into multiple stages, focusing on modularity, user-centered design, real-time interaction, and immersive user experience. Below is a step-by-step explanation of how the intelligent companion system was designed, developed, tested, and deployed.

Needs Analysis and Requirement Gathering

Before development, a detailed requirement gathering phase was conducted. This involved:

- **User Interviews & Surveys:** Understanding the challenges faced by students, job seekers, and casual learners using existing language learning platforms.
- **Persona Development:** Creating learner personas (e.g., college student preparing for interviews, tourist learning travel phrases).
- **Gap Analysis:** Mapping existing apps (Duolingo, Mondly, etc.) against desired features (real-time correction, immersion, gamification).

The findings emphasized the need for:

- A conversational platform beyond scripted answers.
- Real-world usage context.
- Personalized learning that adapts to user proficiency.
- More human-like, interactive learning experiences.

Prototyping and UI/UX Design

Prototyping was done using Canva and Figma for:

- Designing the 3D environment layout (virtual cafe, classroom, meeting room).
- Visualizing user flows: login → level select → interaction → feedback → score.
- Planning the position and animations of the avatar tutor.

UI/UX guidelines ensured:

- Minimalist, intuitive interface.
- Easy navigation, accessible to non-tech-savvy users.
- Responsive design adaptable to various screen sizes.

10.48047/jocaaa.2024.33.08.249

Wireframes and mockups were shared with peers and mentors for feedback before development.

3D Environment and Avatar Development

Using Blender for modeling and Unity/Unreal Engine for development, we built:

- Realistic avatars with facial expressions, gesture animations, and lip-syncing.
- Interactive 3D scenes where learners practice in contextual environments.
- Animation controllers that trigger avatar gestures (nods, smiles, frowns) based on user input and AI feedback.

Key features:

- Modular avatar structure (interchangeable outfits, facial emotions).
- Physics-enabled environment for natural immersion (e.g., virtual object interaction).

Integration of Large Language Models (LLMs)

The backbone of the intelligent companion is its AI-powered conversational engine, using:

- OpenAI GPT-4, Meta LLaMA 3.2, and Groq API.

Key integration steps:

- API endpoints configured to send user inputs and receive contextually relevant responses.
- Backend pre-processing ensures inputs are clean, tokenized, and mapped to learner context.
- Post-processing adjusts response tone and injects avatar emotions.

Personalization achieved by:

- Logging user learning history (vocabulary mastered, grammatical errors).
- Adjusting complexity in real-time (e.g., simpler sentences for beginners).

Backend and Interaction Logic

The interaction engine connects the front-end interface with the LLMs and avatar logic:

- **Backend Stack:** Python Flask (APIs), Firebase (learner profiles and analytics).
- **Conversation Engine:**
  - User speech → STT → Sent to LLM → Response generated → Rendered by avatar using TTS and gestures.
  - Error detection (tense, word choice) highlighted with tips and explanations.
- **Session Management:**
  - Each user has a unique session with saved progress.
  - Adaptive lessons generated using pre-tagged content based on CEFR levels (A1–C2).

Gamification and Reward System

Gamification adds motivation through:

- Scenarios like "Book a hotel," "Give a job interview," or "Debate a topic."
- XP Points for correct sentence construction, fluency, or confidence.
- Streaks & Badges to encourage daily practice and milestone achievements.
- Unlockable Levels: Learners must pass checkpoint quizzes to progress.
- Visual rewards: avatar claps, trophies, and progress bars on the dashboard.

Testing and Evaluation

Multiple layers of testing were performed:

10.48047/jocaaa.2024.33.08.249

- **Functional Testing:** Ensured avatar correctly speaks and gestures according to LLM responses; validated grammar correction, sentence formation, and feedback clarity.
- **UI Testing:** Checked for consistency, button behavior, screen adaptability, and form validation.
- **Device Compatibility Testing:** Tested on different screen resolutions and OS platforms.
- **User Testing:** Feedback collected from 10 student users on ease of use, engagement, and educational value.
- **Performance Testing:** Measured API latency, avatar rendering speed, and scene loading time.

Result: Average load time under 2.5 seconds; 90% of users reported improved fluency confidence after 5 sessions.

Deployment and Accessibility

Deployment options included:

- WebGL build via Unity deployed on GitHub Pages for public access.
- Firebase Authentication for account-based learning progress.
- Planned mobile deployment (Android APK) using Unity's mobile export tools.

Maintenance and Future Scaling

Scalability and maintainability built into the architecture:

a) **Modular Codebase:** Swap avatars, add new languages, or integrate AI models without rewriting core logic.
b) **Update Cycle:** GitHub version control enables continuous integration and feature tracking.
c) **Extensibility:** Future support for multi-user collaboration, classroom mode, or VR learning.

Results

The core objectives laid out during the planning phase of the intelligent companion were successfully met. The system now provides a fully immersive, AI-powered language learning experience. Once implemented, the system is visually represented in Figure 2.

Fig. 2. Intelligent Companion Result

Feature Status Summary

The system integrates several core capabilities essential for an AI-powered personalized learning companion:

- **3D Avatar Tutor:** Fully implemented, enabling realistic and interactive sessions with learners.
- **Real-Time Conversation using LLMs:** Functional, allowing dynamic and context-aware responses.

10.48047/jocaaa.2024.33.08.249

- **Personalized Learning Progression:** Currently under development, ensuring adaptation to individual learner proficiency.
- **Gamified Immersive Scenarios:** Developed to enhance engagement and motivation.
- **User Progress Tracking and Analytics:** Enabled for continuous monitoring and evaluation.

System Performance Summary

The system demonstrates strong performance across multiple dimensions:

- **Page Load Time:** Averages around 2 seconds on standard broadband, aided by an optimized Unity WebGL build.
- **LLM API Response Latency:** Approximately 1.1 seconds per prompt, providing near real-time interaction.
- **Speech Recognition Accuracy:** 92%, with testing focused on Indian English accents.
- **Grammar Feedback Consistency:** 96%, effectively covering common language patterns.

User Testing Feedback

A pilot study was conducted with 10 student participants, each engaging in five distinct learning scenarios, such as ordering food or participating in an interview simulation. The average ratings and user satisfaction results are summarized in Table 2.

Table 2. User Feedback Ratings

Feedback Metric	Avg. Rating (Out of 5)
Ease of Use	4.6
Usefulness of Feedback	4.8
Realism of 3D Interaction	4.5
Engagement Factor	4.9
Overall Satisfaction	4.7

5. Key Observations

User feedback indicated that participants felt more confident when speaking in a structured virtual setting. The avatar's gestures and expressions added emotional depth and realism to interactions. LLM-generated feedback was perceived as natural and supportive rather than robotic or static. Additionally, participants appreciated the progress dashboard and reward system, which they reported as motivating and engaging.

6. Future Scope

While the intelligent companion in its current version demonstrates the successful integration of AI-driven conversation, 3D environments, and adaptive learning, there remains ample room for enhancement and expansion. This version is already quite advanced, but new features could further improve usability and learning outcomes. Currently, interaction with the avatar is primarily via typing. Future improvements could include instant responses through speech input. Key future developments include:

- Multi-Language Support
- Speech-Input and Real-Time Voice Interaction
- Emotion Detection and Adaptive Responses
- Enhanced Gamification and Learning Rewards
- Teacher Dashboard and Learning Pattern Analytics
- AI-Driven Feedback Dashboard

10.48047/jocaaa.2024.33.08.249

7. Conclusion

The implementation of the intelligent companion marks a significant advancement in educational technology. By combining large language models, real-time speech recognition, and immersive 3D environments, it provides interaction and engagement beyond traditional learning applications. It transforms language learning from a passive, text-based task into an active, realistic, and emotionally engaging experience. Moreover, its modular structure ensures scalability, easy integration of future updates, and support for multiple languages. The intelligent companion represents a working prototype of the future of AI-assisted learning.

References

[1] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in *The Adaptive Web*, Berlin, Heidelberg: Springer, 2007, pp. 3–53.

[2] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.

[3] Y. Heffernan and C. Heffernan, "The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching," *Int. J. Artif. Intell. Educ.*, vol. 24, no. 4, pp. 470–497, 2014.

[4] M. Meurers and C.-E. Cheng, "Advantages of natural language processing and computer-assisted language learning," in *Second Language Learning and Teaching with Technology*, 2nd ed., Routledge, 2018.

[5] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," *User Model. User-Adapted Interact.*, vol. 4, pp. 253–278, 1994.

[6] K. VanLehn, "The behavior of tutoring systems," *Int. J. Artif. Intell. Educ.*, vol. 16, pp. 227–265, 2006.

[7] E. Leon, L. T. Sue, and N. Zelinsky-Wibbelt, "AI-based adaptive vocabulary learning for L2 learners," *Comput. Assist. Lang. Learn.*, vol. 30, no. 7, pp. 685–709, 2017.

[8] L. Johnson, S. Smith, R. Willis, A. Levine, and K. Haywood, "The 2011 Horizon Report: K–12 Edition," *New Media Consortium*, 2011.

[9] J. M. Holmes and G. Richards, "Dialog-based intelligent tutoring for language learning," in *Proc. 13th Int. Conf. Artif. Intell. Educ. (AIED)*, Beijing, 2007, pp. 202–209.

[10] L. T. Habegger and K. Karimi, "Chatbots as conversation partners in L2 speaking practice: An empirical study," *Lang. Learn. Technol.*, vol. 23, no. 2, pp. 169–188, 2019.

[11] M. D. Pardos and N. B. Heffernan, "Modeling individualization in a machine learning architecture to improve student learning," in *Proc. 21st Int. Conf. User Model., Adaptation, and Personalization (UMAP)*, Rome, 2013, pp. 315–327.

[12] X. Chen, Z. Qiu, and C. Huang, "Personalized practice scheduling for second language acquisition," *Comput. Educ.*, vol. 125, pp. 396–409, 2018.

[13] S. Li, J. Zhang, and Y. Zhang, "Adaptive recommendation of reading materials in language learning systems," *Educ. Technol. Res. Dev.*, vol. 67, pp. 311–329, 2019.

[14] A. K. Boelmann and C. Conrad, "Designing intelligent language learning companions: A user-centered approach," in *HCI International*, Orlando, 2020, pp. 120–130.

[15] D. Mitrovic, J. Ohlsson, and K. Barrow, "Constraint-based tutors: Past and future," *IEEE Trans. Learn. Technol.*, vol. 1, no. 4, pp. 270–283, 2008.

[16] Y. Xu and R. Warschauer, "Computer-assisted oral language practice with automatic speech recognition: A review," *CALICO J.*, vol. 28, no. 3, pp. 744–760, 2011.

10.48047/jocaaa.2024.33.08.249

[17] H. Godwin-Jones, “Emerging technologies: Mobile-assisted language learning,” *Lang. Learn. Technol.*, vol. 21, no. 2, pp. 2–11, 2017.

[18] J. Heffron and C. Harry, “Adaptive error correction feedback in AI-driven language tutoring,” in *Proc. 2022 Int. Conf. Educ. Data Mining (EDM)*, Durham, NH, 2022, pp. 480–490.

[19] G. Kukulska-Hulme, “Mobile-assisted language learning [special research topic],” *Lang. Learn. Technol.*, vol. 20, no. 3, pp. 1–10, 2016.

[20] S. Lan and P. Dilling, “Toward personalized AI companions for learner engagement: A design study in language education,” *Int. J. AI Educ.*, vol. 31, pp. 102–125, 2021.

[21] R. Pajak, J. Xu, and M. Graesser, “Use of conversational agents to foster L2 vocabulary acquisition: A longitudinal study,” *J. Educ. Psychol.*, vol. 113, no. 4, pp. 739–755, 2021.

[22] B. Smith, X. Jiang, and R. Peters, “The effectiveness of Duolingo in developing receptive and productive language knowledge and proficiency,” *Lang. Learn. Technol.*, vol. 28, no. 1, pp. 1–26, 2024.

[23] Anonymous, “The impact of MondlyVR on language learning,” *Int. J. Educ. Technol.*, 2023.

[24] Anggraini, “Improving students’ pronunciation skill using ELSA Speak application,” *J. English Lang. Educ. Technol.*, vol. 5, no. 1, pp. 135–141, 2022.

[25] A. F. Young, “Top 10 Best AI Language Learning Apps 2025,” 2025. [Online]. Available: https://blog.alexanderfyoung.com/top-10-best-ai-language-learning-apps/