An AI-Driven VR Learning Framework Using RL-Optimized Transformer Models for Personalized English Proficiency Assessment

A. Sri Lakshmi[1], Dr. E.S. Sharmila Sigamany[2], Dr. Revati Ramrao Rautrao[3], Ms. K. Ezhilmathi[4], Dr. Bhuvaneswari Pagidipati[5], Elangovan Muniyandy[6], Dr. Adlin Sheeba[7]
Research Scholar, Department of English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India [1]
Associate Professor, Department of English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India [2]
Associate Professor, Department of Management, Dr. D. Y. Patil B-School, Pune, Maharashtra, India[3]
Assistant Professor, Department of English, Sri Sairam Engineering College, Chennai, India[4]
Associate Professor of English (Ratified by JNTU K), Dept. of English and Foreign Languages, Sagi Rama Krishnam Raju Engineering College (A), Bhimavaram – 534204, West Godavari Dt, Andhra Pradesh, India[5]
Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai - 602 105, India[6]
Professor, Department of Artificial Intelligence and Machine Learning, St. Joseph's College of Engineering, OMR, Chennai, India[7]

***Abstract*—Effective English language learning demands adaptive, interactive, and flexible instructional support, which traditional e-learning systems and existing AI tutors struggle to provide due to limited immersion, static feedback mechanisms, isolated task structures, and the absence of robust reward-driven learning strategies. Although prior studies on VR-based learning environments and Natural Language Processing (NLP) have reported enhanced learner motivation and engagement, most existing solutions suffer from fixed task sequencing, limited real-time linguistic intelligence, and inadequate grammar and pronunciation correction capabilities. To address these challenges, this study proposes a Virtual Reality–based architecture named the Self-Evolving Neural Intelligence Tutor (SENIT), driven by Curriculum Reinforcement Learning and Hierarchical Adaptive Weighting. SENIT integrates a fine-tuned T5 transformer for grammar refinement and prosody-aware feedback, while a reinforcement learning agent dynamically adjusts task difficulty and lesson progression based on learner performance. Developed using Python and TensorFlow and deployed within a Unity3D VR environment, SENIT enables realistic conversational simulations and multimodal learner assessment. Experimental evaluation on a dedicated VR English Learning Dataset demonstrates grammar and pronunciation accuracy improvements of 90% and 81%, respectively, outperforming existing models by approximately 12 percentage points. Additionally, learners achieved notable fluency gains and high engagement scores, highlighting SENIT's effectiveness in delivering personalized, immersive language learning experiences.**

Keywords*—AI-driven learning; Virtual Reality; English language education; reinforcement learning; Natural Language Processing

I. Introduction

The fast development of artificial intelligence and immersive technologies has brought a major shift in online education, especially in the fields of second language learning [1] [2]. Learning the English language needs continuous interaction and immersion into the situations and adaptive feedback systems that respond to the needs of individual learners [3]. Nevertheless, the traditional e-learning platforms and most of the current AI-based tutoring systems are still limited to a fixed content delivery, a lack of contextual awareness, and rule-based feedback models [4] [5]. Such systems are likely to be based on predetermined learning routes and individualized practices, which limit the involvement of learners and do not consider the differences in the level of proficiency, pronunciation, and the cognitive learning pace [6] [7]. Even though recent progress in Natural Language Processing and intelligent tutoring systems have enhanced automated assessment, the absence of immersive conditions and the dynamic adjustment of pedagogy still remains a hindrance to their success in achieving real-world communicative competence [8] [9].

Virtual Reality (VR) has become one of the solutions of the future, as it provides realistic and interactive learning environments that can improve the motivation of learners and experiential learning [10] [11]. In spite of this aspect, majority of VR-based learning system of English provide non-adaptive series of tasks, limited real-time linguistic intelligence and lack of grammar-pronunciation correction system and reward-driven individualization mechanisms [12]. In order to fill these gaps, this study proposes Self-Evolving Neural Intelligence Tutor (SENIT), a VR-based learning platform that operates on Curriculum Reinforcement Learning and Hierarchical Adaptive Weighting. SENIT combines a refined Transformer model to refine grammar and provide feedback with a learning agent that adapts task difficulty, order of the lessons and feedback techniques dynamically to the performance of a learner. The framework used in this study combines Python and TensorFlow with Unity3D to provide immersive conversational simulations and individualized assessments with the goal of significantly enhancing fluency, accuracy, and engagement of learners in VR-based systems of learning English language.

A. Research Motivation

This study is motivated by the rising discrepancy between the dynamic needs of an effective English language learning and the limited flexibility of the current digital learning systems [13]. Although conventional e-learning tools and the existing AI tutors assist in automated learning and testing, they mostly employ the one-dimensional delivery of the material, a set of predetermined routes, and general feedback, which cannot consider the unique characteristics of learners in terms of their proficiency, pronunciation and learning speed [14]. Despite the potential of Virtual Reality based learning environments to increase engagement and motivation in learning through immersion, the majority of the current VR solutions do not include real-time linguistic intelligence, adaptive task sequencing, and reward-driven personalization [15]. Besides, the lack of built-in grammar and prosody correction tools inhibits the capacity of supporting authentic conversational competence in them. The given restrictions outline the necessity and the development of an intelligent, self-adaptive learning framework that integrates both immersive VR experiences and sophisticated Natural Language Processing and reinforcement learning. As a result, this project is inspired to develop a VR-permitted, AI-enhanced tutoring platform that dynamically personalizes teaching, feedback, and evaluation, making it possible to allow a learner to continue their development, become more fluent and remain engaged in learning the English language.

B. Significance of the Study

This research is important as it contributes to further development of personalized and immersive learning of languages, which is achieved through the combination of Virtual Reality, Transformer-based linguistic intelligence, and the use of Adaptation (also created with the help of Reinforcement Learning) into a single framework [16]. The SENIT architecture suggested is more flexible than the current VR learning systems because it provides curriculum-conscious sequencing of tasks, adaptive prioritization that is hierarchical and reward-based and allows the instruction to dynamically adapt to the performance of the learners. The system can be used to achieve more authentic and context-dependent language evaluation than existing methods because of the addition of fine-tuned Transformer models to perform real-time grammar refinement and prosody-sensitive pronunciation feedback. The practical effectiveness of the presented framework is underlined by the empirical evidence of the improvement in grammar accuracy, pronunciation performance, fluency, and learner engagement. Academically, this work offers a scalable and extensible framework to intelligent tutoring systems, whereas educationally, it offers a plausible avenue of offering highly immersive and personalized English language learning that can be transferred to virtual classes, adult training, and distance learning.

C. Problem Statement

Traditional language learning systems are not able to offer the individual, dynamic, and interactive interaction, particularly in an immersive experience, such as VR. Even the existing AI solutions are too biased towards one-task-based approaches (primarily grammar-fixing) and fail to address such crucial aspects of the problem as prosody, the process of interaction between the learner and the system, and task sequence with pedagogical significance. Moreover, most systems lack difficulty adjustment within real-time and this leads to improper correspondence with the actual progress of learners and normally leads to a deterioration in motivation and performance [17]. The VR-based language tutors also rarely include the Reinforcement Learning (RL) in the curriculum-based system; this does not allow them to propose guided and progressive learning paths [18]. To overcome these weaknesses, this study will suggest a hybrid adaptive system, which will combine grammar correction, prosodic analysis and RL-based curriculum development to sustain active development, further engagement and reactive learning of the English language.

D. Key Contributions

- Presents the SENIT, an English learning architecture based on VR, which personalizes the instruction dynamically based on Curriculum RL.
- Introduces an agent based on RL whose adaptive weighting is hierarchical to change the difficulty of lessons, the choice of tasks and the learning rate of the learner depending on his real-time performance.
- Includes a refined T5 Transformer of real-time grammar correction and prosody-sensitive pronunciation assistance in immersive VR dialogues.
- Integrates VR simulations with Python and TensorFlow to use Unity3D as a basic but powerful tool that supports conversational practice and multimodal learner assessment.
- Shows strong enhancement of grammar accuracy, pronunciation, and fluency and learner engagement as opposed to current state-of-the-art models.

E. Rest of the Study

The rest of the study is structured as follows: Section II presents some of the relevant studies on AI and VR-based language learning. Section III gives the materials and methods, namely the integration of T5-based NLP, RL, and speech analysis inside the MVL-RE VR learning environment. Section IV mentions the results and performance evaluation. Lastly, Section V is the conclusions drawn from this study and future study directions.

II. Literature Review

Belda-Medina and Marrahi-Gomez [19] suggested that augmented reality in English as a second language instruction is an innovative help toward addressing most of the age-old problems characterized by traditional language learning. Traditional teaching by itself does not engage students into the real application of the language; therefore, their motivation declines and their pace of learning slows down. Otherwise, with AR, it incorporates digital content into the real environment, thus supporting the enhancement of learning on the basis of real interaction in a real-world experience, in making learning contextualized, instant feedback, and allowing the student to acquire language skills in a very engaging manner. Various ways in which AR can be used to teach some of the most

important aspects of the English language: vocabulary, pronunciation, grammar, reading, and writing. By superimposing AR elements such as digital labels, grammatical clues, and interactive text on real-world objects, students can participate in enjoyable and meaningful hands-on practice. Finally, AR offers students the opportunity to practice their language skills in simulated environments and real-world contexts, thus aiding retention and application of their learning. The extra layer that AR provides to ESL instruction could catalyze the improvement of existing pedagogies and methods, rendering the learning of languages effective and enjoyable for non-native speakers.

Chen et al. [20] proposed that acquisition of foreign languages must be through active engagement with input, while meaningful task fulfillment follows. PBL gives real-life context and hands-on experiences, while VR interaction occurs in multimodal informational settings. The extent to which the inclusion of VR technologies into PBL enhances students' motivation, problem-solving skills, and vocabulary while learning EFL. A sample of 84 engineering students enrolled in an English for Specific Purposes course was assigned randomly either to an experimental group or to a control group. In the experimental group, students participated in a VR-enhanced PBL environment in which they explored problem scenarios via VR and produced their own VR videos that proposed solutions for the aforementioned problems. In turn, the control group worked on the very same problem scenarios in a traditional PBL setting that did not integrate VR. After the intervention, all participants engaged in problem-solving analysis, vocabulary knowledge test, motivation for learning questionnaire, and individual interviews. It was found that students under VR-assisted PBL exhibited increased vocabulary acquisition and motivation in learning English regarding future careers, while no significant difference was noticed in problem-solving performance between the two groups. These results underline the possible advantages of immersive environments of VR in promoting contextualized learning in EFL education. Nonetheless, some limitations of the study must be acknowledged in interpreting the findings, and future studies should follow to probe further into these issues to facilitate the effective integration of VR in language learning.

Lu [21] incorporates VR, which will be interactive, in English language learning and will be used in immersive environments to reinforce engagement, communication skills, and speaking experience. It uses the RHH-ISVM method to measure interaction among the students by studying patterns of communication, linguistic complexity, and learning development in VR conditions. Following noise filtering and Min -Max normalization, PCA ultimately provides the crucial features, including fluency, richness of vocabulary, and textual suitability. The framework then makes learning individualized through the customization of instructional material based on these features that have been extracted, and in this way, they allow adaptive feedback and focal enhancement. The VR and intelligent ML modelling have a high potential of providing tailored and successful language learning experiences. The study, however, admits the constraints of sample representativeness and data collection biases that ought to be dealt with in future research.

Raman et al. [22] suggested VR-integrated lessons on students' engagement, motivation, and autonomous learning of ESL students and developing their verbal communication skills. Factors between these three variables, in a VR-based learning environment, were also explored. This mixed-method study involved a sample of senior students-majority of the respondents from the total sampling of 166 high school students. Survey data were analyzed through SPSS and uncovered that these largely engaged (M=4.55), motivated (M=4.36), and autonomously learned (M=3.90) participants were, in fact, high scorers. Correlation analysis also proved that the engagement, motivation, and autonomy of learners in VR-based lessons are strongly related. It was observed that the learners have extremely positive attitudes towards VR-based language learning. Thematic analysis revealed that the following benefits included realistic experiences, more retention, more involvement, and better communication skills. VR promotes verbal and digital literacy and makes learners ready to live in the world of technology. Nonetheless, the definite changes of long-term benefits of English communication under the conditions of constant VR consumption require the conduct of larger research involving a wider sample of participants.

Li et al. [23] state ESP cuts across several issues. These include vocabulary understanding and retention-the major challenges faced by learners. A crucial aspect of ESP teaching focuses on enhancing vocabulary acquisition and practical use. An innovative and emerging audiovisual tool, VR, has implemented this kind of incidental vocabulary learning by stressing presence and immersion. The experiential property affords students the opportunity to learn vocabulary more effectively, thus promoting the learning outcome and participation. This identified how far the students acquired vocabulary relating to the hydrologic cycle. The participants consisted of 36 geography students divided between experimental and control groups, with an experimental group of 18 learners using VR-based learning technology and an equal number in the control group learning through the traditional video-based method. Findings indicated that those in the VR-assisted learning way are better incidental vocabulary learners and definitely higher on cognitive, behavioral, and social engagements than those in the control groups. However, limitations have been recognized; VR resources were limited in that the setting for vocabulary learning was not entirely immersive throughout the process in ESP contexts.

Xie et al. [24] stated that globalization is rapidly advancing technological processes, human interaction, and social behavior. These international skills require the need to market such individuals. This model is aimed at combining the immersion-based English instruction and VR technology so that language acquisition can be improved. The quasi-experiment involved 106 pupils in one of the Chinese schools within a 16-week time span, the sessions lasted three hours a week and added upto 48 hours. Statistical software was used to test the proposed hypotheses with the help of the collected data of the experiment. These results showed that there was a high positive correlation between VR and immersion-based language teaching and immersion-based teaching and academic achievement. Moreover, the VR had a positive correlation with learning outcomes. In contrast to the conventional study, this

breakthrough also established another test analysis of oral test advancement to enhance learning. Depending on the results, suggestions are drawn to further the English teaching and the evolution of linguistics. However, limited by a rather limited sample size and the limited time of the experiment, which calls into question the generalizability of its conclusions, because the results are very promising.

Fang-Ying Riva Lo [25] enhances the communication skills of EFL learners to bring them to the point where they actually achieve authentic language acquisition. Traditional textbook-based instruction does not offer that kind of excitement or engagement with language; thus, it does not help the real growth in language ability. This high-immersion VR, on the other hand, transforms the whole picture of self-directed learning experiences within simulated real-world environments into a greater role for engagement and participation by learners. This calls for a 3D learning environment named VRLE, developed to give an authentic learning ecosystem for communicative skills development. Hence, the 72 low-achieving junior-high-level students participated. It was mostly quantitative data combined with qualitative data that would determine the adequacy of VRLE. It consisted of pre- and post-tests for communicative performance, a questionnaire known as the IPQ for determining presence perception by the students, and semi-structured interviews. The great contributions of VRLE for improved English communication skills and the sense of presence for students in the learning environment. The participants also showed favorable attitudes toward the VR-based learning experience. The findings support the argument about the importance of high-immersion VR in EFL education. Other accessibility-related areas still cannot be exhausted within the upcoming investigation to optimize the inclusion of VR into language learning. This adds to the ever-increasing pile on VR for language education, but much more still needs to be done in exploring its applicability for communication skills development among low-achieving EFL learners.

According to the current literature, VR and AR have proven to be able to aid ESL/EFL learning positively by enhancing vocabulary acquisition, pronunciation, engagement, and contextual communication by using immersive environments. AR also aids practice and immediate feedback whereas DL-based methods, including RHH-ISVM used in combination with project-based learning have demonstrated an improvement in motivation and performance. Nevertheless, the majority of VR-based systems are not dynamic, and they do not feature real-time linguistic intelligence, adaptive control of instruction, and closed loops of AI-guided learning. The previous AI systems consider single-task processes without considering the multimodal feedback, reinforcement-based curriculum planning, or affective engagement modelling. Limited size of samples and brief durations of interventions also pose limitations to generalizability, and more sophisticated, scalable, intelligent VR systems are required.

III. Framework for AI-Driven Personalization and Immersive Language Learning

The suggested approach is based on a modular end-to-end design comprising immersive VR, adaptive, and assessed with sophisticated AI. The first step is interaction of learners in a Unity3D VR setting that emulates real-life scenarios of conversation and records multimodal interactions, such as speech, textual reactions, and interaction behavior. A finely-tuned T5 Transformer model processes these inputs to do real-time grammar correction and prosody-sensitive pronunciation analysis. At the same time, a Curriculum Reinforcement Learning agent monitors learners' performance measures, such as accuracy, fluency, response time, and engagement, and calculates adaptive rewards in a hierarchical weighting approach. According to the experienced policy, the RL agent will dynamically rearrange the sequence of tasks, difficulty of the lessons and intensity of feedback to align with the developing competence of the learner. The whole system is deployed with Python and TensorFlow to train and infer the model, and the smooth integration with the VR interface allows providing constant feedback and feedback. This is a closed-loop adaptive learning process that enables the SENIT to facilitate individualized, immersive, and increasingly challenging instruction in the English language. The proposed method workflow is shown in Fig. 1.

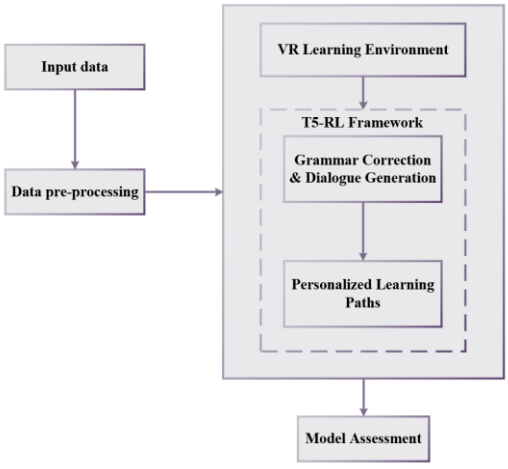

Fig. 1. Workflow of the proposed approach.

A. Data Collection

The Kaggle Impact of Virtual Reality on Education dataset is an aggregation of quantitative and qualitative data that was collected in educational institutions where VR technologies were deployed to examine the effect VR has on learning outcomes and experiences. It can be used to assess the proposed English learning system since the features of its engagement, interaction, and the use of VR directly reflect the involvement and immersion of the learner, which is complemented by the measures of linguistic performance, recorded in the tailor-made VR English Learning Dataset. It generally consists of student-related features, including demographic variables, engagement, creativity, and academic achievement, and includes indicators of the application and impact of VR in instructional settings, e.g., the relationship between indicators of the improvement of VR immersive use on courses and levels. It is possible to correlate VR use with such educational achievement indicators as engagement, comprehension, creativity, or other measures of learning success using the dataset, allowing the researcher to understand the transformational power of VR in learning and teaching settings [26].

B. Data Pre-Processing

In data preprocessing, the raw data gathered in learning through VR-based education interactions will be converted into a model-ready and structured form in order to determine reliability and consistency in learning evaluation. Firstly, missing values in numerical variables (engagement score, time to complete the task, or performance ratings) are processed with the help of statistical imputation methods, either mean or median imputation, denoted as Eq. (1):

$$x_i^* = \begin{cases} x_i, & if\ x_i \neq NaN \\ \frac{1}{N}\sum_{j=1}^{N} x_j, & if\ x_i = NaN \end{cases} \quad (1)$$

where, x_i^* is the imputed value, and N is the number of valid observations. Thereafter, a feature normalization is conducted to remove scale variations among variables, including pronunciation accuracy, fluency score and engage level. Min-max normalization is used to distribute features to a single range [0, 1], which is as shown below Eq. (2):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

where, X is the original value, $X_{max} and\ X_{min}$ are the maximum and minimum values in the dataset and X' is the normalized value. The step avoids dominance of the high magnitude features when training the model. Lastly, the cleaned and normalized data is formatted into learner state vectors $S_t = [g_t, p_t, f_t, e_t]$, which are grammar accuracy, pronunciation score, fluency, and engagement at time t. These hierarchical representations match the aims of the datasets for English proficiency assessment and allow the effortless combination with Transformer-based linguistic analysis and reinforcement learning-based customization.

C. AI Model Framework

The two elements are essential to support the structure of the proposed AI model, as they are synergistic to one another to create an intelligent, adaptive, and real-time interactive VR-based environment to learn English. The former involves the initial part that involves fine-tuning a T5 Transformer model to make the finer grammar correction and dialogue generation to bring about context-sensitive and more natural language interaction, resembling more realistic communication scenarios. The second component encompasses an RL agent that will personalize the learning experience based on the performance and the degree of engagement, and the achievement of an individual learner. This agent keeps varying the difficulty of the lesson, the order in which the tasks and feedback are presented, and delivers the best learning outcomes. A combination of these modules forms a self-evolving ecosystem where AI and VR intersect to make the training of the English language very immersive, responsive and effective based on the requirements of the specific learner.

1) Fine-tuning T5 for grammar correction and dialogue generation: Google developed the T5 model, a pre-trained NLP transformer that can be fine-tuned for the multi-task NLP problems such as grammar correction and dialogue generation. For this study, T5 is fine-tuned with supervised learning over a dataset of corrected and uncorrected English sentences to learn the patterns of correct grammar and sentence structures.

a) Fine-tuning process: Given a dataset D containing pairs of incorrect and corrected sentences, as expressed in Eq. (3):

$$D = \{(X_i, Y_i)\}_{i=1}^{N} \quad (3)$$

where, X_i represents an incorrect sentence (input), Y_i represents the grammatically corrected sentence (target output), and N is the total number of sentence pairs in the dataset. The T5 model is trained to minimize the cross-entropy loss between the predicted sequence $\acute{Y}_i$ and the target sentence Y_i, using the objective function, as given in Eq. (4):

$$L = -\sum_{i=1}^{N}\sum_{t=1}^{T} Y_{i,t} log P(\acute{Y}_{i,t} | X_i, \theta) \quad (4)$$

where, L represents total loss, N represents data samples, $\acute{Y}_{i,t}$ represents predicted label, and $Y_{i,t}$ represents true label.

2) Implementing RL for personalized learning paths: The RL agent will undertake to adapt the difficulty level of lessons, customize exercises, and ensure that it is intended for optimal progression depending on the learning gathered from the student's output performance. This is achieved under a Markov Decision Process framework. It is given in Eq. (5):

$$MDP = (S, A, P, R, \gamma) \quad (5)$$

where, S denotes the set of states, A denotes set of actions, P represents the probability function, R denotes reward function, and γ denotes discount factor. In order to guarantee progress by learners, reward function $R(s, a)$ promotes improvements in three important learning aspects. It is given in Eq. (6):

$$R(s, a) = w_1 \cdot \Delta F + w_2 \cdot \Delta G + w_3 \cdot E \quad (6)$$

where, w_1, w_2 and w_3 are weight factors. E denotes engagement metrics, ΔF denotes change in fluency score, and ΔG denotes change in grammar score. The Q learning or DQN

are utilized by the RL agent that selects learning actions that will result in maximizing cumulative rewards for each learner. The Bellman equation is defined as the Q value update equation. It is calculated using Eq. (7):

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left(R(s,a) + \gamma \max_{a'} Q(s',a') - Q(s,a)\right) \quad (7)$$

where, α represents learning rate, $R(s,a)$ represents immediate reward, and $\max_{a'} Q(s',a')$ represents the maximum Q-value. With this approach, learners receive real-time, adaptive, and personalized feedback, and progress at this pace so that they improve at their own ability. With the system continuously evaluating a learner's response and keeping exercises adjusted accordingly, AI-driven VR language learning is more effective than static, one-fits-all solutions. The process of the personalized learning path is given in Fig. 2.

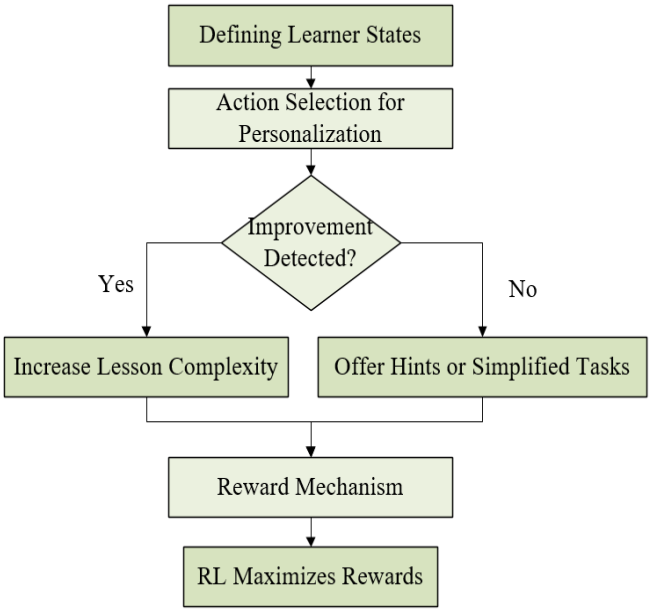

Fig. 2. Process of personalized learning path.

D. Integration of T5 and RL in VR-Based Learning

NLP grammar correction and dialogue generation using T5 and enhanced by RL controlled learning path adaptation system constitutes the foundation of a highly interactive, intelligent, and personalized VR-based learning of English. The T5 model corrects the grammar and refines the sentences, making the learners get the real-time feedback of what they are saying or writing in English. T5 is opposed to the Language learning tools of all the tools, essentially learn with preset exercises. However, unlike this, T5 is critical of what is being typed by a learner, underlines the errors in grammar and consequently provides more subtle suggestions of how a mistake-free sentence can be typed in the stipulated time. This will allow the learners to correct themselves in real-time and learn the proper language structure, and hasten the process of proficiency. Students also get exposed to the free, unscripted, fluent dialogues rather than the pre-scripted dialogues, hence becoming more fluent and comfortable when speaking English.

1) Grammar correction function: The grammar correction feature is based on the T5 model to convert input provided by the learner into grammatically modified text. It detects structural, syntactic and contextual errors, and creates polished sentences, which facilitates language enhancement in real-time, fluency and immediate corrective cues in the VR learning environment. It is measured in Eq. (8):

$$\hat{y} = T5(x) \quad (8)$$

It represents the T5 model transforming learner input x into corrected output $\hat{y}$, enabling real-time grammar refinement and improved sentence structure during VR-based language learning interactions.

2) RL state and action selection in adaptive learning: The RL element represents the proficiency, engagement, and accuracy of each learner as states and chooses the actions that best maximize policy, like by modifying the difficulty or type of task. This will facilitate individualized, interactive learning adjustment to optimize learning growth and long-term involvement in real-time. It is prescribed in Eq. (9):

$$s_{t+1} = f(s_t, a_t) \quad (9)$$

where, the models show how the learner's next state is determined by current proficiency, engagement, and the RL-chosen action, supporting real-time adaptation of lesson difficulty within the VR environment.

E. VR Learning Environment

VR learning environment offers a simulative and interactive platform, and learners communicate in a life-like and dialogue situation with AI-based virtual characters. It mimics real-life examples such as job interviews, small talks, boarding trains, hotel enquires and business meetings that allow the learners to apply the English language in real-life. In this type of simulation using AI, the environment is dynamic to react to the reaction of the learner, and fluency, proper use of language, and confidence in speaking English are promoted. The system also helps in the effective acquisition of the language due to the safe, involving, and contextual environment, better communication, and the preparedness of the learners to work with English in real-life scenarios. It is measured in Eq. (10):

$$r = g(u, c) \quad (10)$$

where, the function maps learner utterance u and conversational context $c$to an AI-generated response r, enabling natural, unscripted VR dialogues and context-aware feedback during real-time interactions.

1) Developing realistic conversation scenarios: To establish an efficient and dynamic learning experience, the system will include a large range of conversational situations based on learners on the beginner and intermediate levels. The learning conditions will dynamically change with learning experience, where an RL-based learning-assist mechanism will tailor the experience based on the performance of the different

learners and the contextual needs of the learner in real-time. Virtual characters will be created with the help of AI and will be based on NLP and sentiment analysis, which will enable them to track the level of learner engagement, such as the level of questions being asked and interactions initiated by the learner. These personalities will offer real-time, context-sensitive feedback on pronunciation, grammatical, or conversation errors, either with a natural human-like voice reply, or with suggestions of possible corrections where feasible

2) AI-driven interaction and feedback system: The VR-based Language Learning System provides a dynamic refinement of the learner's proficiency through the smart, real-time multimodal feedback mechanism incorporating auditory, textual, and visual stimuli. It is an NLP-based speech recognition and spellcheck application that uses the T5 NLP model to examine both spoken and written text, identify grammar and pronunciation mistakes, and make successful corrections in specific contexts. Phonetic guidance is facilitated by speech synthesis and accurate forms of language are solidified through the repetition of practice. The virtue of AI-driven virtual characters, the conversations would be unscripted and natural and the pre-programmed dialogues would be replaced with realistic and adaptive conversations. The adaptive RL-based system is a new paradigm that also adapts the lesson difficulty to the performance of the learner to provide individual and efficient learning processes. Combining the grammar fixing of T5, lesson personalization powered by RL, and the presence of immersive AI interactions, the platform overcomes the barriers of the traditional approaches to language teaching. Such a combined method will promote fluency, accuracy, interest, and confidence, creating a more real, responsive and efficient context in learning the English language. It is illustrated in Eq. (11):

$$F = \phi(e_g, e_p, e_s) \tag{11}$$

where, the feedback module processes grammatical errors e_g, pronunciation deviations e_p, and semantic issues e_s, generating precise multimodal corrective signals to refine learner fluency and linguistic accuracy.

F. Curriculum RL with Multimodal Reward

In order to enhance flexibility, we incorporate a curriculum-based learning plan into the reinforcement learning agent. The curriculum planner increases the level of difficulty of the work in the next stages (beginner-intermediate-advanced) that promotes the gradual development of the learners. At all the levels, the agent maximizes the multimodal reward comprising of: 1) grammar improvement through the T5 correction module, 2) prosodic speech quality (intonation and rhythm), and 3) engagement cues (frequency of interaction and time-on-task). The reward function is a combination of the two functions. In the proposed reinforcement learning framework, the reward r_t is designed to reflect the learner's overall progress at each interaction step. It is computed as a weighted sum of improvements in grammar accuracy (Δg_t), pronunciation accuracy (Δp_t), fluency gain (Δf_t), and the current engagement level (e_t), as given by Eq. (12):

$$r_t = \alpha.\Delta g_t + \beta.\Delta p_t + \gamma.\Delta f_t + \delta e_t \tag{12}$$

where, $\alpha, \beta, \gamma, \delta$ represent adaptive weights which determine the relative significance of each factor. This reward is given immediately following every learning interaction, which will enable the agent to keep on varying the task difficulty and sequence. The design is such that the instructional choices are informed by linguistic enhancement and by the interest of the learner, which underpins a tailored learning and gradual learning process.

G. Hierarchical Curriculum with Adaptive Weighting

Hierarchy Curriculum with Adaptive weighting (HCAW) builds upon the foundation of Curriculum RL and provides a dynamic ranking of the sequence of tasks and rebalancing of multimodal rewards based on the changing proficiency of individual learners. However, a different approach, contrary to the traditional linear curriculum, is HCAW, which is designed as a hierarchical graph, with each node modeling clusters of skills, e.g., pronunciation, grammar, or prosody, and the edges model the prerequisite relation between skills. A meta-controller constantly reviews the learner status and performance metrics in order to choose the following suitable task node and dynamically change the reward weights. The general learning task is to get the best expected payoff and encourage lean, customized and streamlined competencies learning in every learner. It is described in Eq. (13):

$$J(\theta, \emptyset) = E_{\pi\theta, \mu\emptyset}\left[\sum_{t=0}^{T} \gamma^t r_t\right] \tag{13}$$

where, $\pi\theta$ is the student policy and $\mu\emptyset$ represent the meta controller. The empirical outcomes of HCAW are better convergence of the policy network and scoring higher than static curricula, which is why it is a successful fit to the VR-based language learning environment.

Algorithm 1. Adaptive Hierarchical Curriculum RL for VR English Learning

```
Input: Learner state s_t (proficiency, grammar, pronunciation, engagement)
Output: Selected task h_l, updated reward weights α_t, β_t, γ_t

Initialize policy π_θ, meta-controller μ_φ
Initialize α_0, β_0, γ_0
Initialize task hierarchy H={h_1,...,h_L}
for each episode:
   Observe current state s_t
   if learner improvement in grammar < threshold_g:
      select task h_l related to grammar
   else if learner improvement in pronunciation < threshold_p:
      select task h_l related to pronunciation
   else:
      select next hierarchical task h_l in H
   Compute reward r_t = α_t Δg_t + β_t Δp_t + γ_t e_t
   Update α_t, β_t, γ_t based on moving averages
   Update policy π_θ using Q-learning/DQN
end for
```

The HCAW Hierarchy in VR English learning algorithm is discussed in the pseudocode provided in Algorithm 1. It takes

the condition of a learner, his or her proficiency, grammar, pronunciation and engagement as the inputs and takes the decision of task and the new weight of reward as the outputs. The meta-controller compares the performance depending on the state of the if-else conditions: in case of grammar improvement poor, the grammar tasks are chosen; in case of pronunciation improvement poor, the pronunciation tasks are chosen; otherwise, the next task in the hierarchy is chosen. Multimodal rewards are computed and weight modification is done depending on the advance of the learner. Finally, the RL agent modifies its policy by using Q-learning or DQN, which ensures that the learning process is more customized, adaptive and progressive.

IV. Results and Discussion

The proposed AI-VR language learning system, which is written and implemented in Python, has demonstrated considerable improvement in other areas of learning. The system has a fine-tuned T5 grammar correction model, an adaptive sequencing, RL agent, and AI-generated conversations that enable a learner to experience a highly responsive learning process. Results indicate that grammatical, pronunciation and adaptive curriculum-based improvement were greatly improved. The study has shown that the level of immersion among learners was increased and enabled by the feedback loop based on dynamically customized task structures and the real-time feedback loop. The natural conversation interactions were made possible through the virtual world, and fluency and contextual interpretation were improved. Overall, the findings prove that AI-based personalization in VR is a potent stimulator of linguistic performance that offers an intelligent, interactive, and adaptable medium by which an individual can acquire knowledge about the English language in a successful way.

TABLE I. Simulation Parameters and Hardware Setup

Parameter	Specification
Programming Environment	Python 3.10, TensorFlow 2.12, PyTorch 2.1, Unity 2022 with XR Toolkit
NLP Model Used	Fine-Tuned T5 Transformer (base, 220M parameters)
RL Algorithm	DQN with ε-greedy policy and $\gamma = 0.9$
VR Device Configuration	Oculus Quest 2, 90Hz refresh rate, 1832x1920 resolution per eye
GPU Used for Training	NVIDIA RTX 3090 (24 GB VRAM, CUDA 11.8, cuDNN 8.6)
System Memory	64 GB DDR4 RAM
Audio Processing	Real-time Speech-to-Text via Google Speech API + Noise Suppression
Dataset Split Ratio	80% Training, 10% Validation, 10% Testing (Total Samples: ~30,000)

Table I is a simulation environment, which involves the application of high-performance GPU acceleration and real-time audio processing and continuous RL in means of a streamlined DL architecture operating in the VR environment. All the parts are set to the best of their capabilities and adjusted to the optimum configuration to obtain the performance of the specific experimentation setup. The system offers efficient training of models, and it also gives immersive and high-resolution VR experiences with low-latency interaction to accommodate reactive and individualistic learning. This system provides scalability and flexibility in order to attain dynamic and effective acquisition of the English language with the AI-VR framework proposed.

TABLE II. Grammar and Pronunciation Score Improvements

Metric	Pre-Intervention (Mean ± SD)	Post-Intervention (Mean ± SD)	Improvement
Grammar (%)	55 ± 4	90 ± 3.5	+35
Pronunciation (%)	50 ± 3.8	81 ± 3.2	+31

Table II presents the average grammar and pronunciation scores and standard deviation of the means before and after the process of AI-VR intervention. Grammar also improved greatly as the point result has improved by +35 to 90 +- 3.5% and pronouncing improved by +31 to 81 +- 3.2. The standard deviations are also quite small that implies consistency in the performance improvement across the participants. The results confirm the effectiveness of Curriculum Reinforcement Learning using Hierarchical Adaptive Weighting in an immersive VR environment and that it can offer learners substantial and steady gain in mastery of both English grammar and pronunciation.

A. Performance Outcomes

Fig. 3 outlines the progress of the learners throughout 10 accommodations of VR instruction of the English language. The grammar, pronunciation and engagement scores were significantly better with an increase of 55 to 90, 50 to 81 and 60 to 82 respectively and this indicated the development of learning. It can be seen that the scores are improved in every episode with scores that are not overlapped by the markers. As shown in this graph, the Curriculum RL with Hierarchical Adaptive Weighting has the benefit of improving the performance of the learners progressively, between grammar, pronunciation, and engagement. The diagram highlights the real-time adjustment and demonstrates the usefulness of the AI-based VR system of language learning.

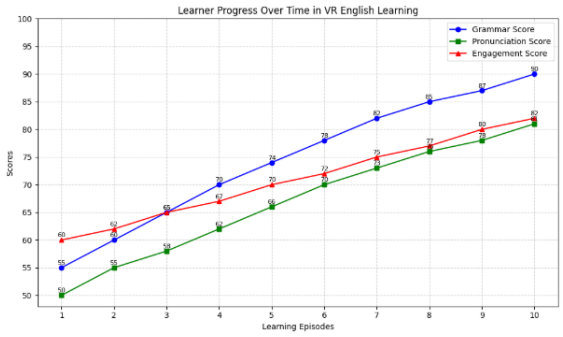

Fig. 3. Learner progress over time in VR English learning.

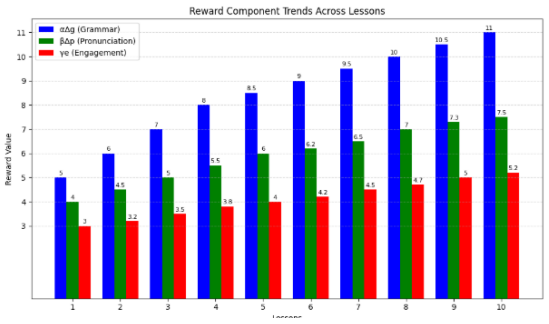

Fig. 4. Reward component trends across lessons.

Fig. 4 shows trends of reward components through the 10 lessons in the VR English learning system. Grammar reward (Δg) was improved 5 to 11, pronunciation reward (2 2) was improved 4 to 7.5 and engagement reward (3 2) was improved 3 to 5.2, which indicates gradual focus on each of the learning dimensions. It is clear and there is no overlapping since there are annotated values above bars. The figure demonstrates that Hierarchical Curriculum with Adaptive Weighting could dynamically modify rewards based on the performance of the learner to balance between the improvements of grammar, the improvement of pronunciation, and to engage the learner, thus, adaptive, multimodal reinforcement in the AI-driven VR learning environment.

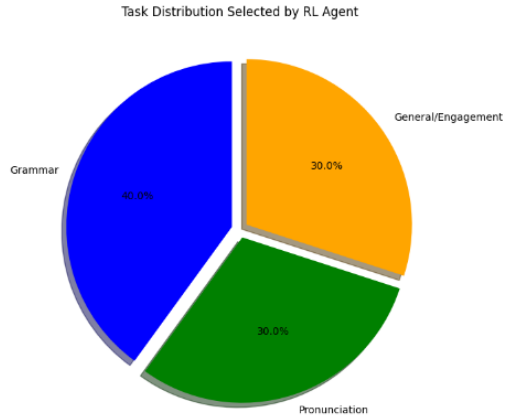

Fig. 5. Task distribution selected by the RL agent.

Fig. 5 presents the distributions of task selection of the RL agent in the VR English learning system. Out of all the tasks that are chosen during the lessons, Grammar-oriented tasks take 40% of the area, Pronunciation tasks take 30% and General/Engagement tasks take 30 percent of the area, which is in line with the priorities of the agent that varies depending on the performance of the learners. These percentages demonstrate that the Hierarchical Curriculum, where Adaptive weighting is used, allows the RL agent to dynamically choose the tasks on which it needs to work so that it might devote the same amount of attention to grammar, pronunciation, and engagement. The visual metaphor is a reference to the personalized and adaptable nature of the AI-enhanced VR learning.

Fig. 6 shows the steady increase in the training and validation accuracy, which depicts the increasing capacity of the system to forecast learner performance accurately. The close gap between curves proves the high generalization rate, which testifies the success of the reinforcement-based curriculum sequencing and multimodal feedback on the VR language learning background.

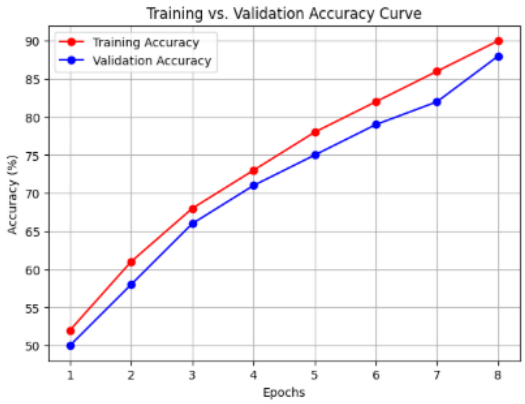

Fig. 6. Training and validation accuracy.

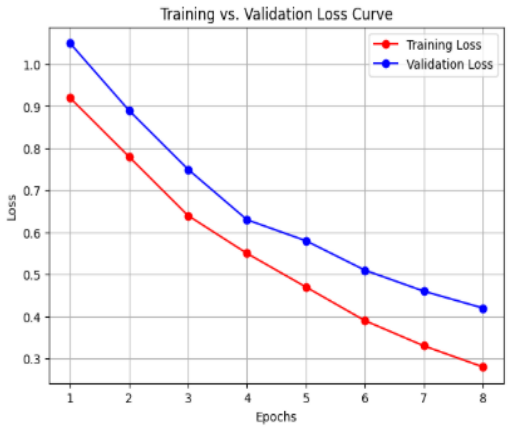

Fig. 7. Training and validation loss.

Fig. 7 illustrates a steady decrease in training and validation loss in the epochs, which means the convergence of the model. The two curves conforming to each other in a close sense imply that there is good generalization, effective parameter learning and sound optimization in an AI-VR adaptive language learning system.

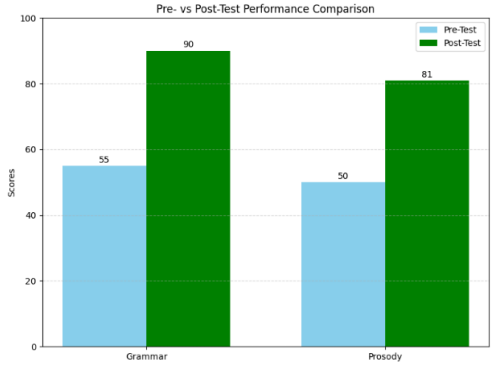

Fig. 8. Pre- versus post-test performance comparison.

The performance regarding the VR English learning environment is compared in Fig. 8 before and after the test. Grammar points were increased to 55-90, as well as the prosody

points which were increased to 50-81, with a heavy improvement of the scores of the learners who were exposed to the AI-based VR system. Annotated values above the bars show the gains without overlapping. This visualization demonstrates Curriculum RL with Hierarchical Adaptive Weighting is successful since learners succeeded not only at the grammar but had skills in the pronunciation. The figure illustrates the individualistic and adaptable pedagogical impact of language interventions founded on VR.

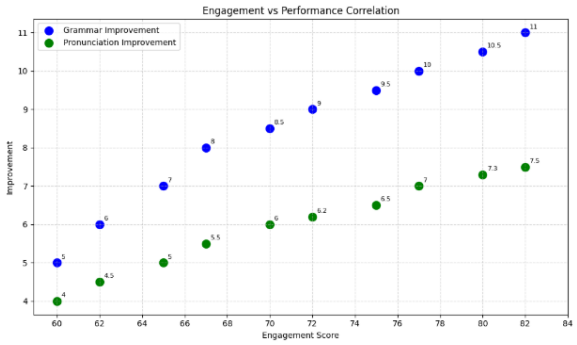

Fig. 9. Engagement vs. Performance correlation.

The relation between the performance and the engagement of the learners in 10 episodes is revealed in Fig. 9. The gains in grammar were measured at as high as 5 to 11 points and in pronunciation as high as 4 to 7.5 points which is accompanied by the gains in engagement that range at 60 to 82. The points of data at individual level are separated clearly to avoid overlaps. The visualization was correlated with the engagement and it was discovered that the more engaged the people the better they become in grammar and pronunciation. These results confirm that Curriculum Reinforcement Learning with a Hierarchical Adaptive Weighting is effective and support the fact that AI-VR system is an adaptive and interaction-based language learning system that enhances the performance of language learning.

B. Performance Metrics

This section statistically assesses the effectiveness of the suggested AI-VR framework in enriching English language learning. Core performance metrics are quantified to determine improvements in accuracy of grammar, fluency, engagement, responsiveness to feedback, and adaptive learning. The established formulas standardize the assessment to allow for objective analysis of learner performance and system efficacy.

TABLE III. PERFORMANCE METRICS OF AI-DRIVEN VR ENGLISH LEARNING SYSTEM

Metric	Grammar Accuracy (%)	Pronunciation Accuracy (%)	Fluency Improvement (ΔF)	Prosody Improvement (ΔP)	Engagement Score (E)
Pre-Test Mean	55	50	0	0	60
Post-Test Mean	90	81	5–11	4–7.5	60–82
Improvement / Gain	+35	+31	+5–11	+4–7.5	+0–22
Standard Deviation	4	3.5	1.5	1.2	6
Max Value	95	85	11	7.5	82
Min Value	55	50	5	4	60

Table III and Fig. 10 gives a summary of the performance measures of learners in VR English learning environment. The grammar strength increased by 55% to 90, pronunciation by 50 to 81, fluency gains were between 5 and 11 points, prosody between 4 and 7.5 points and engagement increased to between 60 and 82. Standard deviations show variability of learners. High and low values indicate the best and the worst performers. It is shown in this table that the Curriculum RL with Hierarchical Adaptive Weighting, which includes T5 grammar correction, RL-driven adaptive lessons, and immersive VR interactions, is effective in achieving significant learning improvements in several language skills, which are measurable.

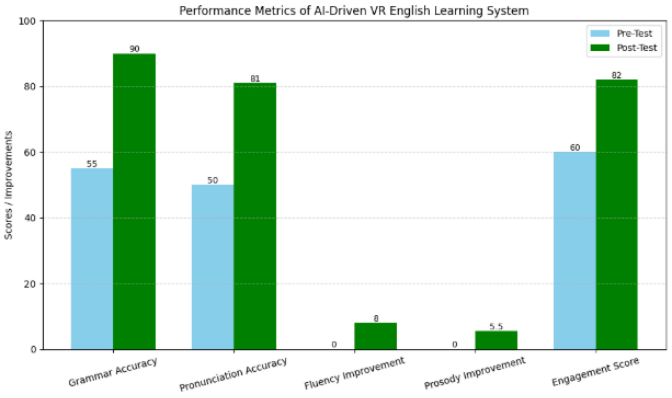

Fig. 10. Performance metrics of the proposed AI-VR method.

TABLE IV. PERFORMANCE COMPARISON WITH EXISTING METHODS

Model / Method	Grammar Accuracy (%)	Pronunciation / Prosody (%)	Fluency Improvement (ΔF)	Engagement Score (E)
DMBO-BiGRU [27]	78	70	4	65
RHH-ISVM [28]	72	68	3	62
CNN-LSTM [29]	80	73	5	67
Proposed VR AI System	90	81	8	82

Table IV provides the comparison of the Proposed VR-AI system with three of the existing approaches: DMBO-BiGRU, RHH-ISVM, and CNN-LSTM. Accuracy of grammar in the proposed system is up to 90, which is better compared to DMBO-BiGRU (78%), RHH-ISVM (72%), and CNN-LSTM (80%). Pronunciation/prosody is also enhanced to a significant level of 81 as compared to the current methods of 70-73. The improvement of fluency is 8 points, which is greater than 3-5 points of other models. The engagement score is 82, which presents higher participation of learners. This is a clear indication that the VR-AI system is the best in language acquisition, pronunciation, fluency, and engagement of the learner.

C. Discussion

The combination of Curriculum reinforcement learning and Hierarchy Adaptive weighting within a virtual learning environment is a very effective model in acquiring English language. The significant changes in grammar accuracy (55 + 4 to 90 + 3.5) and pronunciation performance (50 + 3.8 to 81 + 3.2) prove that the proposed system allows providing specific, adaptive training based on constant performance monitoring changes. The fact that fluency has increased by eight points and the engagement score is relatively high at 82 also suggests that the idea of immersive interaction and task advancement that is based on a reward greatly contributes to the motivation of the learner and the consistency of learning. The excellent results of the proposed approach as compared to DMBO-BiGRU, RHH-ISVM, and CNN-LSTM baselines confirm the fact that adaptive task sequencing and multimodal, context-aware feedback are the key points of enhancing learning. Furthermore, the association between engagement and performance is positive, which implies that immersive VR-based interaction facilitates deeper thought involvement and long-term knowledge acquisition. The fact that the system can dynamically develop, evolve in real-time and gives feedback in contexts also highlights its scalability and use in areas other than language learning especially in competency-based training and skills development. Although this study has promising results, there are some limitations. The test was conducted in a controlled VR setting using a small sample of data, which might not reflect a variety of real-life learning settings. Transformer and reinforcement learning model use impose computation costs, which may be a constraint when used on low-resource systems. VR equipment and audio quality also may influence the accuracy of pronunciation. In addition, long term learning retention has not been studied extensively, which should be investigated in the future.

V. Conclusion and Future Work

The Curriculum RL when combined with Hierarchical Adaptive weighting in a learning environment through VR proves more effective in teaching the English language, as it provides adaptive, individualized, and more immersive learning. The overall outcomes of the SENIT, grammar accuracy, pronunciation, fluency, and engagement with the learners are strong and above the state-of-the-art base models. Transformer-based linguistic intelligence along with real-time feedback and adaptive task sequencing are effective in supporting individual paths of learning, encouraging learners and enhancing long-term memory. These findings support the idea that context-based, immersive and dynamically changing VR learning environments present a solid ground on which effective language acquisition and the development of interactive skills may be achieved.

Future studies can focus on the further development of the SENIT framework to facilitate learning in more than two languages by adding more linguistic characteristics like semantics, pragmatics, and discourse-level comprehension. Emotion recognition and affective computing would also be helpful in improving engagement and personalized feedback. The mass implementation in actual classroom or distance learning environments will permit testing the scaling, long-term learning retention and cross-cultural efficiency. Also, VR combined with augmented reality (AR) or mixed reality (MR) environments may be even more immersive and more effective in context learning, and most importantly, more computationally efficient, which will allow its use in the form of edges to make it more accessible to all and adapt to real-time in different learning settings.

References

[1] R. Lege and E. Bonner, “Virtual reality in education: The promise, progress, and challenge,” JALTCALL, vol. 16, no. 3, pp. 167–180, Dec. 2020, doi: 10.29140/jaltcall.v16n3.388.

[2] E. Analyti et al., “Virtual Reality in Education: Transforming Learning through Immersive Technology,” Technium Education and Humanities, vol. 10, pp. 1–11, Oct. 2024, doi: 10.47577/teh.v10i.11766.

[3] “(PDF) The Use of Virtual Reality (VR) on English Student’s Speaking Skills in Educational Context: A Systematic Literature Review,” ResearchGate, Oct. 2024, doi: 10.30998/jedu.v4i2.11325.

[4] L. Wei, “Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning,” Front. Psychol., vol. 14, Nov. 2023, doi: 10.3389/fpsyg.2023.1261955.

[5] W. Yan, V. L. Lowell, and L. Yang, “Developing English language learners’ speaking skills through applying a situated learning approach in VR-enhanced learning experiences,” Virtual Reality, vol. 28, no. 4, p. 167, Nov. 2024, doi: 10.1007/s10055-024-01061-5.

[6] “(PDF) A COMPARATIVE ANALYSIS OF TASK-BASED AND TRADITIONAL INSTRUCTION IN ENGLISH LANGUAGE ACQUISITION,” ResearchGate, Oct. 2024, Accessed: Feb. 21, 2025. [Online]. Available: https://www.researchgate.net/publication/380786041_A_COMPARATIVE_ANALYSIS_OF_TASK-BASED_AND_TRADITIONAL_INSTRUCTION_IN_ENGLISH_LANGUAGE_ACQUISITION

[7] Falsk Raza, “AI in Education: Personalized Learning and Adaptive Assessment,” 2023, Unpublished. doi: 10.13140/RG.2.2.24796.77446.

[8] “(PDF) The Impact of Virtual Reality Technology on Student Engagement and Learning Outcomes in Higher Education,” in ResearchGate, Accessed: Feb. 21, 2025. [Online]. Available: https://www.researchgate.net/publication/374910266_The_Impact_of_Virtual_Reality_Technology_on_Student_Engagement_and_Learning_Outcomes_in_Higher_Education

[9] A. Parmaxi, “Virtual reality in language learning: a systematic review and implications for research and practice,” Interactive Learning Environments, vol. 31, no. 1, pp. 172–184, Jan. 2023, doi: 10.1080/10494820.2020.1765392.

[10] M. S. Anwar et al., “Immersive Learning and AR/ VR-Based Education,” in Cybersecurity Management in Education Technologies, 1st ed., New York: CRC Press, 2023, pp. 1–22. doi: 10.1201/9781003369042-1.

[11] O. E. Cinar, K. Rafferty, D. Cutting, and H. Wang, “AI-Powered VR for Enhanced Learning Compared to Traditional Methods,” Electronics, vol. 13, no. 23, p. 4787, Dec. 2024, doi: 10.3390/electronics13234787.

[12] G. Sarsar, Djameleddine Derias, and Y. Karaca, “AI-Powered Procedural Content Generation: Enhancing NPC Behaviour for an Immersive Gaming Experience,” 2023, doi: 10.13140/RG.2.2.23719.52647.

[13] C. Papakostas, “Artificial Intelligence in Religious Education: Ethical, Pedagogical, and Theological Perspectives,” Religions, vol. 16, no. 5, p. 563, 2025.

[14] M. El Hajji, T. Ait Baha, A. Berka, H. Ait Nacer, H. El Aouifi, and Y. Es-Saady, “An Architecture for Intelligent Tutoring in Virtual Reality: Integrating LLMs and Multimodal Interaction for Immersive Learning,” Information, vol. 16, no. 7, p. 556, 2025.

[15] W. Wiboolyasarin, K. Wiboolyasarin, P. Tiranant, N. Jinowat, and P. Boonyakitanont, “AI-driven chatbots in second language education: A systematic review of their efficacy and pedagogical implications,” Ampersand, vol. 14, p. 100224, 2025.

[16] O. E. Cinar, K. Rafferty, D. Cutting, and H. Wang, “AI-Powered VR for Enhanced Learning Compared to Traditional Methods,” Electronics, vol. 13, no. 23, Art. no. 23, Jan. 2024, doi: 10.3390/electronics13234787.

[17] M. Aly, "Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model," Multimedia Tools and Applications, vol. 84, no. 13, pp. 12575–12614, 2025.

[18] C. Song, S.-Y. Shin, and K.-S. Shin, "Optimizing foreign language learning in virtual reality: a comprehensive theoretical framework based on constructivism and cognitive load theory (VR-CCL)," Applied Sciences, vol. 13, no. 23, p. 12557, 2023.

[19] J. Belda-Medina and V. Marrahi-Gomez, "The Impact of Augmented Reality (AR) on Vocabulary Acquisition and Student Motivation," Electronics, vol. 12, no. 3, p. 749, Feb. 2023, doi: 10.3390/electronics12030749.

[20] C. Chen, H. Hung, and H. Yeh, "Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners," Computer Assisted Learning, vol. 37, no. 3, pp. 851–860, June 2021, doi: 10.1111/jcal.12528.

[21] S. Lu, "Interactive virtual reality environments for English language learning: Immersive technologies to enhance language acquisition and communication skills," Journal of Computational Methods in Sciences and Engineering, p. 14727978251318803, Feb. 2025, doi: 10.1177/14727978251318803.

[22] K. Raman, H. Hashim, and H. H. Ismail, "Enhancing English Verbal Communication Skills through Virtual Reality: A Study on Engagement, Motivation, and Autonomy among English as a Second Language Learners," IJLTER, vol. 22, no. 12, pp. 237–261, Dec. 2023, doi: 10.26803/ijlter.22.12.12.

[23] Y. Li, S. Ying, Q. Chen, and J. Guan, "An Experiential Learning-Based Virtual Reality Approach to Foster Students' Vocabulary Acquisition and Learning Engagement in English for Geography," Sustainability, vol. 14, no. 22, Art. no. 22, Jan. 2022, doi: 10.3390/su142215359.

[24] Y. Xie, Y. Liu, F. Zhang, and P. Zhou, "Virtual reality-integrated immersion-based teaching to English language learning outcome," Frontiers in psychology, vol. 12, p. 767363, 2022.

[25] F.-C. O. Y. Fang-Ying Riva Lo, "Facilitating Communicative Ability of EFL Learners via High-Immersion Virtual Reality," 2024.

[26] "Impact of Virtual Reality on Education 🥽 👓." Accessed: Feb. 21, 2025. [Online]. Available: https://www.kaggle.com/datasets/waqi786/impact-of-virtual-reality-on-education

[27] Y. Lou, "The impact of virtual reality environments on English language acquisition: Innovative immersive learning technologies for communication skills development," Journal of Computational Methods in Sciences and Engineering, p. 14727978251337950, 2025.

[28] S. Lu, "Interactive virtual reality environments for English language learning: Immersive technologies to enhance language acquisition and communication skills," Journal of Computational Methods in Sciences and Engineering, p. 14727978251318803, 2025.

[29] T. Su and L. P. Rue, "The analysis of English language teaching with machine translation based on virtual reality technology," Scientific Reports, vol. 15, no. 1, p. 15845, 2025.