Adaptive Learning Companions: Enhancing Education with Biosignal-Driven Digital Human

Alaeddin Nassani
University of Aizu
Aizuwakamatsu, Japan
alaeddin@u-aizu.ac.jp

John Blake
University of Aizu
Aizuwakamatsu, Japan
jblake@u-aizu.ac.jp

Julian Villegas
University of Aizu
Aizuwakamatsu, Japan
julian@u-aizu.ac.jp

Figure 1: Implementation of the biosignal-adaptive learning system: (left) The digital human tutor interface as rendered in the Unity-based VR environment, displaying a ConvAI-powered interaction; (right) A participant equipped with Quest 3 and the Samsung Watch 7 for real-time biosignal adaptation for language learning.

Abstract

We introduce a novel approach to language learning leveraging digital humans as adaptive tutors within immersive XR environments. Our system's novelty lies in the use of biosignals, specifically real-time heart rate data, collected from a Samsung Watch 7, to dynamically adapt the learning experience. The digital human tutor adjusts its behavior, feedback, and the difficulty of the learning content based on the learner's inferred cognitive and emotional state. We present the fully developed system architecture, which integrates a customizable digital human powered by ConvAI, LLM, an XR environments, and a data streaming pipeline. While human participant testing is planned , preliminary insights from the system's development demonstrate the technical feasibility of this approach. This research has the potential to significantly enhance language learning outcomes, engagement, and motivation by creating more personalized, and engaging learning experiences, paving the way for a new generation of adaptive educational technologies.

CCS Concepts

• **Human-centered computing** → **Virtual reality**.

Keywords

Virtual Reality, Language Learning, Adaptive, Biosignal, Avatar, Digital Human

ACM Reference Format:
Alaeddin Nassani, John Blake, and Julian Villegas. 2025. Adaptive Learning Companions: Enhancing Education with Biosignal-Driven Digital Human. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3706599.3719877

ACM ISBN 979-8-4007-1395-8/25/04
https://doi.org/10.1145/3706599.3719877

1 Introduction

Immersive language learning, particularly in Extended Reality (XR) environments, offers exciting possibilities, including interactive practice with virtual characters and exposure to nonverbal communication cues. However, a challenge lies in replicating the dynamism and responsiveness of real-world interactions. A critical

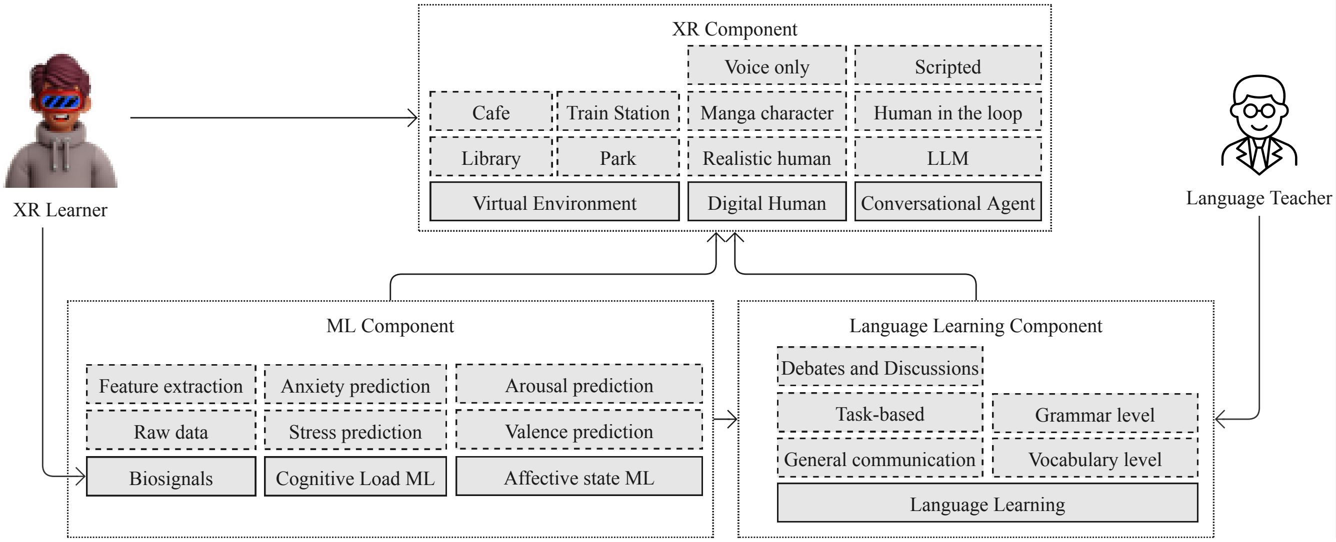

Figure 2: System architecture of the Adaptive Learning System (ALS), featuring three core components: Extended Reality (XR) for immersive interaction, Machine Learning (ML) for personalized adaptations, and a Language Learning component for targeted educational support.

limitation in these systems is the current lack of adaptation to learners' physiological and cognitive states. Unlike interactions with human tutors, who naturally adjust to emotional and cognitive cues by reading non-verbal signals such as facial expressions, body language, and gestures, as well as vocal tone and speech patterns, most XR language learning environments remain static, offering a one-size-fits-all approach regardless of the learner's engagement, frustration, or cognitive load. However, it should be noted that humans vary in their ability to accurately interpret such signals, and cross-cultural differences can lead to misinterpretation. For example, an Indian student shaking their head side-to-side might be mistakenly perceived as a sign of disagreement, when in fact it often conveys understanding or agreement in Indian culture. Similarly, Japanese learners may smile or laugh when feeling nervous or unsure, which could be misinterpreted by Western English speakers as amusement or lack of seriousness, rather than a polite response to discomfort or uncertainty.

This research builds on theories of empathy and emotion regulation, recognizing their critical role in learning and social interaction. Rodrigues et al.'s [23] process model of empathy provides a framework for designing virtual agents capable of realistic empathic interactions, modulated by factors related to both agent and user. Furthermore, theories of emotional contagion [11] and emotional mimicry [12] suggest that synchronized emotional responses are key to fostering empathy and social bonds. Therefore, virtual learning environments that adapt to a learner's emotional state hold significant potential for enhancing engagement and knowledge retention.

Despite these insights, a significant gap exists in their application to language learning within XR environments. While recent studies [10, 24] have explored the potential of biosignal-driven adaptations in XR to improve user experience and interaction quality, these efforts have not fully addressed the unique demands of language acquisition. The application of such adaptive technologies to language learning, particularly in tailoring virtual environments and interactions based on real-time biosignals, remains largely unexplored.

This work aims to fill this research gap by investigating how XR environments can be dynamically tailored to language learners' physiological and cognitive states to create more effective and engaging learning experiences. By incorporating models of empathy and emotion regulation into the design of a digital human tutor, we aim to develop more adaptive and responsive learning environments that enhance language acquisition and improve the overall learning process. We investigate the following question: **RQ** How can we design a biosignal-aware digital human tutor in an XR environment that adapts to a learner's physiological and cognitive states to enhance language acquisition?

2 Related Work

This research builds upon a growing body of work exploring the intersection of immersive technologies, adaptive learning systems, and language education.

2.1 Adaptive Learning Systems

Spoken Dialogue Systems (SDS) have significant potential to support language learning by acting as virtual interlocutors, providing students with opportunities to practice speaking in interactive and engaging ways. Their potential is further amplified by integrating adaptive systems that tailor learning content based on biosignal data from learners, enabling even more personalized and effective educational experiences. SDS have been effectively utilized for various purposes, including scenario-based practice [27], role-plays [28], and task-based interactions [19]. Furthermore, the integration of large language model (LLM)-based technologies has enhanced SDS

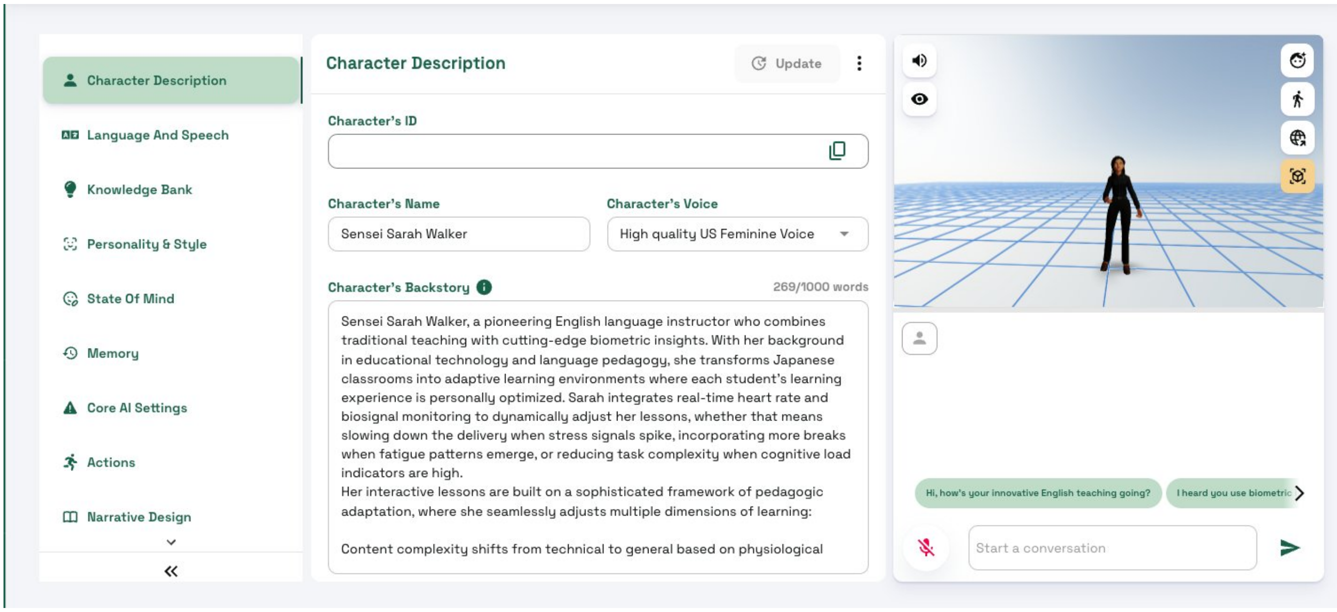

Figure 3: LLM-based avatar using Convai

capabilities for both the development and assessment of speaking skills, as demonstrated in studies exploring automated speaking evaluations and interaction analysis [6, 8, 18, 27]. These advancements underscore the role of SDS in fostering effective language learning experiences. However, SDS can be enhanced by integrating adaptive learning systems (ALS), allowing them to dynamically adjust interactions and feedback based on individual user progress and learning needs.

ALS aim to personalize the learning experience by tailoring content and pacing to individual student needs. Research in this field has shown promising results, with adaptive systems demonstrating effectiveness in improving learners' language proficiency across various levels [26]. A comprehensive review of the literature on adaptive learning systems used in language learning found that instruction was personalized based on multiple factors, including learner proficiency levels and learner needs [13]. The majority of the systems targeted the learning of English by university-level language learners [13] although other languages have been investigated, e.g., French [3]. Prior studies have demonstrated the potential of VR to enhance language acquisition by creating interactive and immersive learning experiences. For instance, Repetto [22] and Pai et al. [20] highlight the benefits of using intelligent avatars within virtual environments to facilitate language practice and improve conversational skills. While these systems offer some degree of personalization, they often rely on pre-programmed rules or performance metrics, limiting their ability to adapt in real-time to the learner's fluctuating cognitive and emotional states. Schultz & Maedche [25] emphasize the value of interpreting biosignals to tailor educational content and adapt to individual learner needs, leading to improved engagement and outcomes. However, the application of real-time biosignal adaptation has not been fully explored in conjunction with digital human tutors in XR environments for language learning.

2.2 Embodied Conversational Agents in Education

Embodied conversational agents, including digital humans, have shown promise in educational settings by enhancing engagement, motivation, and social presence. In the context of language learning, virtual agents can serve as valuable tools for practicing communication skills and receiving feedback. Research on embodied interaction with adaptive robots has demonstrated positive effects on engagement and learning outcomes in language education [29], suggesting the importance of embodied cognition in the learning process. Furthermore, multi-modal conversational agents that integrate natural language processing with visual modalities have been explored as a means to facilitate personalized language learning through more natural and intuitive interactions [9]. The use of these agents can create a more supportive and engaging learning environment, particularly when they are designed to be socially intelligent and responsive to learners' needs.

2.3 Biosignals in HCI and Learning

Biosignals, such as electroencephalography (EEG), electrocardiography (ECG), and galvanic skin response (GSR), offer a window into users' cognitive and emotional states, providing valuable data for designing adaptive systems. Recently, there has been growing interest in developing biosignal-adaptive virtual avatars to enhance social presence and non-verbal communication in virtual environments. Kim & Hong [16] proposed a method for extracting biological signals from face images to create more lifelike avatars without the need for additional hardware. Lee et al. [17] investigated user perceptions of heart rate and breathing rate visualizations in social VR,

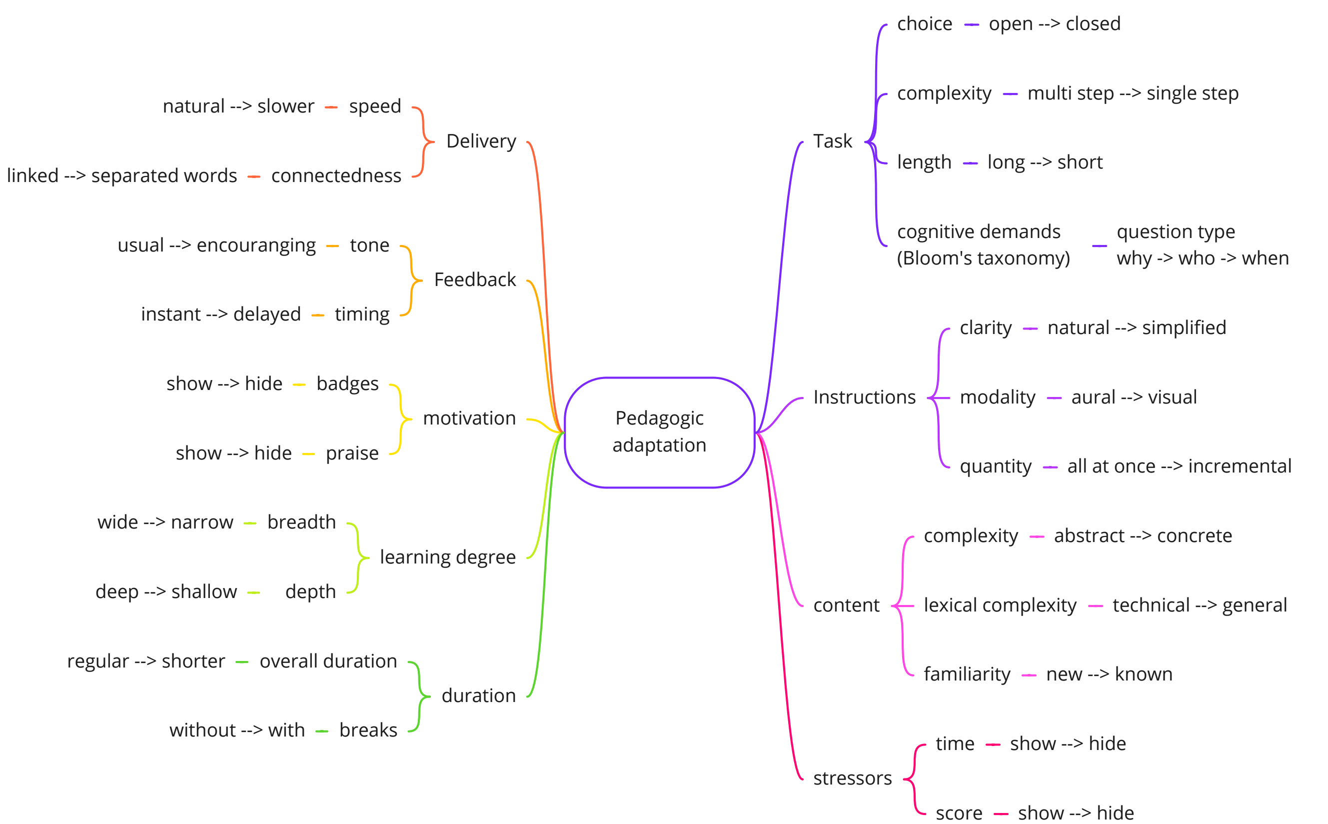

Figure 4: A mind map detailing the parameters of pedagogic adaptation for language learning employed in the system. The map is organized around a central concept of "Pedagogic Adaptation" and branches out into ten categories: Delivery (speed, connectedness), Feedback (tone, timing), Motivation (badges, praise), Task (complexity, length, cognitive demands, choice), Instructions (clarity, modality, quantity), Content (complexity, lexical complexity, familiarity), Stressors (time, score), Learning Degree (breadth, depth), Duration (overall duration, breaks), and Choice (open, closed). Each category includes specific parameters that can be adjusted to personalize the learning experience. For instance, under "Delivery" the "speed" can be adjusted from "natural" to "slower".

finding that designs mimicking real-world objects improved the inference of arousal states while minimizing distractions. Further, studies explored how biosignals can enhance virtual avatars, including through enhanced hand motion reconstruction accuracy using EEG and EMG signals by Fernández-Vargas et al. [5]. These studies demonstrate the potential of biosignal-driven avatars to enrich social interactions in virtual spaces. Collectively, they underscore the potential of immersive technologies and adaptive systems to enhance language learning by tailoring experiences to individual needs, providing engaging and interactive feedback, and ultimately improving motivation and learning outcomes [4, 20].

However, a gap remains in integrating real-time biosignal adaptation with digital human tutors within XR environments for dynamic and personalized language learning. Our work aims to address this gap by developing a system that leverages physiological data to inform the behavior of a digital human tutor and adapt the XR learning environment, fostering a more responsive and effective language learning experience.

3 System Design and Implementation

The system comprises three interconnected components: a Digital Human & XR component responsible for the user interface and interaction, an Adaptive Learning component that dynamically adjusts the learning content based on the user's state, and a Biosignal component that acquires and processes physiological data. Fig. 2 illustrates the overall system architecture

3.1 Digital Human and XR Component

The digital human tutor was developed using Unity (version 2022.3.38f1) and integrated with a customized version of the ConvAI SDK, which manages the avatar's appearance, animation, and AI-driven behaviors. The ConvAI (Fig. 3) setup was tailored using a prompt that defines the avatar's role as an English tutor capable of adapting to physiological input. This prompt included sample English exercises (e.g., "find the error in this sentence") and specified that the avatar should utilize the Claude-3-5-Sonnet[1] LLM model for its underlying

[1] https://www.anthropic.com/news/claude-3-5-sonnet

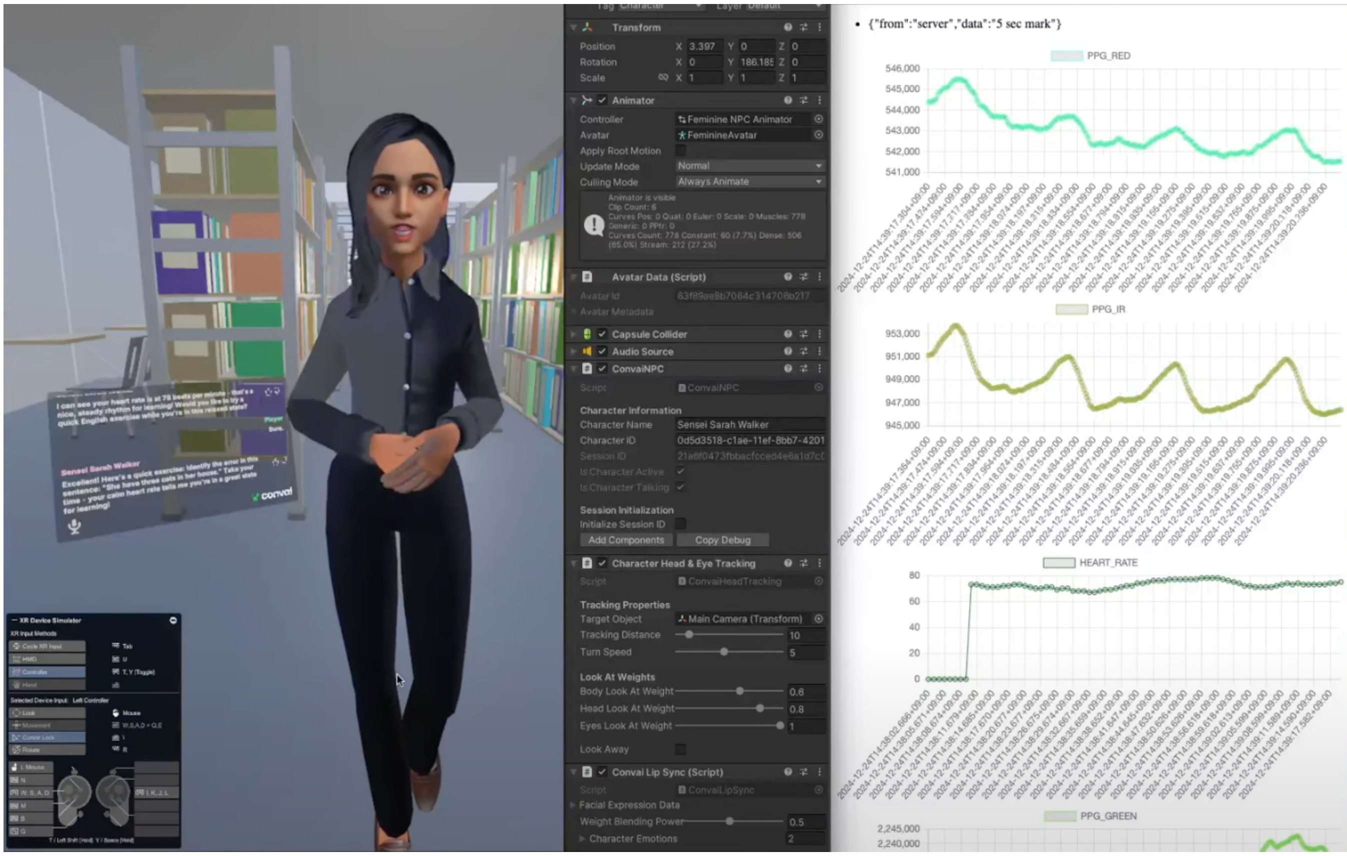

Figure 5: Biosignal (right) raw data PPG data and Average HR streamed to a PC running the avatar on Unity (left).

conversation. The SDK was further modified to receive the user's average heart rate data, streamed at 1 Hz, allowing the avatar to incorporate this physiological information into its adaptive logic. The application was deployed to a Meta Quest 3 headset, rendering an immersive VR scene of a classroom where the digital human, acting as the teacher, stands in front of the user. To enhance engagement and realism, the avatar is programmed to maintain eye contact with the user when they are within a single unit radius (approximately one meter in virtual space) within the Unity environment.

3.2 Adaptive Learning Component

In a study of English language learner perceptions of stressors in traditional classes, Khajavy et al. [14] identified four main stressors, namely lack of preparation, assessment, speaking English, and a lack of understanding.

Stress related to lack of preparation can manifest when, for example, a Japanese student is unexpectedly asked to give a presentation. The anxiety about speaking accurately under-prepared can result in prolonged silences or hesitations. Stress stemming from assessment often emerges before quizzes or exams, as pre-existing anxiety about performance intensifies [2]. For some learners, past negative experiences with language learning further contribute to a fear of failure, heightening stress levels. Speaking English is another major source of stress [21], particularly in social situations. Anxiety about accents, self-consciousness, and fear of negative judgment from peers or instructors can exacerbate feelings of unease and inhibit performance. Finally, stress related to a lack of understanding arises when students struggle to comprehend instructions or classroom tasks [7], hindering their ability to execute activities correctly. This type of stress is common when instructions are delivered quicker than learners are able to process or use unfamiliar language.

An increase in heart rate may be detected in all of these cases, as stress activates the sympathetic nervous system, triggering a fight-or-flight response. If such an increase is detected, a digital human tutor can adapt lesson delivery in real time. For instance, if an elevated heart rate is observed after an instruction is given, the tutor could speak more slowly, paraphrase the instructions using simpler language, or directly ask the student if clarification is needed. As illustrated in Fig. 4, the delivery speed could be adjusted. Instead of relying solely on connected speech, the tutor could repeat a sentence: first with careful enunciation of each word and then more naturally, incorporating the typical linking of words across junctures. This adaptive approach can help alleviate stress and improve learner comprehension, which is in line with humanistic learning theory regarding the creation of a nurturing learning environment [1].

3.3 Biosignal Component

The Biosignal component of the system captures physiological data from the user via a Samsung Watch 7, leveraging its Privileged Health SDK to access raw sensor readings. Specifically, the system receives raw photoplethysmography (PPG) data from infrared (100 Hz), red (100 Hz), and green (25 Hz) sensors, electrocardiography (ECG) data at 500 Hz, and average heart rate (HR) at 1 Hz. To facilitate data transfer, a custom Android applet running on the watch streams the raw sensor data using WiFi to a local computer acting as a WebSocket server. The Unity application, which hosts the digital human and XR environment, subscribes to the HR data stream from this server, enabling real-time updates on the user's average heart rate. Fig. 5 shows an example of raw PPG data and the average HR streamed from the watch to the digital human avatar.

Currently, the system utilizes the average HR to inform the adaptive logic of the digital human tutor. However, future iterations will leverage the full spectrum of raw PPG, ECG, and accelerometer data to extract time and frequency domain features. This will enable the development of a machine learning model for detecting user stress, characterized in terms of valence and arousal, to further enhance the personalization and responsiveness of the learning experience.

4 Preliminary Insights & Discussion

Our development process has demonstrated the technical feasibility of integrating a biosignal-adaptive digital human tutor within an XR learning environment. The real-time data streaming from the Samsung Watch 7 to the Unity application, combined with the digital human's responsiveness to average heart rate, provides a solid foundation for our system's core functionality. However, we acknowledge several challenges associated with wearable sensor data.

Latency in the data pipeline—primarily due to the wireless connection between the watch, phone, and computer, as well as the processing overhead of the SDK on the watch—occasionally affects the system's real-time responsiveness. While this is manageable for initial testing with average heart rate, it may become a limiting factor when integrating higher-frequency data streams, such as raw PPG and ECG. Additionally, the accuracy and reliability of physiological data are susceptible to motion artifacts and individual physiological variations (e.g., skin color, sweat rate), as highlighted in related work by Khan et al. [15].

Currently, our Biosignal Component relies primarily on average heart rate data. Recognizing that a more detailed analysis of the raw signal could yield richer insights into the user's cognitive and emotional states, we are actively investigating signal processing techniques—including both time- and frequency-domains analyses—to extract meaningful features. Furthermore, we are studying the potential interactions between ECG and PPG signals to assess whether their combined analysis can enhance the accuracy and robustness of stress and engagement inference. Although the optimal method for integrating these multiple signals remains under exploration, this ongoing work represents a key direction for future iterations of our system.

We hypothesize that our biosignal-adaptive digital human tutor has significant potential to enhance language learning outcomes, engagement, and motivation compared to traditional methods. By integrating real-time physiological monitoring, the system dynamically adapts lesson delivery based on heart rate data, creating a more personalized and responsive learning environment. This adaptive approach directly addresses our research question by ensuring that learning experiences are tailored to the learner's cognitive and emotional state, fostering more effective and immersive language acquisition.

For instance, the digital human tutor can adjust the pace of instruction, offer encouragement during moments of frustration (as inferred from heart rate patterns), or provide more challenging exercises when the learner exhibits signs of high engagement and low cognitive load. This personalized feedback and scaffolding, informed by continuous physiological monitoring, can potentially optimize the learning process, keeping learners within their zone of proximal development.

The ability to practice language skills in realistic virtual scenarios, coupled with real-time feedback that is sensitive to the learner's emotional and cognitive state, could lead to improved learning outcomes and a more positive and enjoyable learning experience overall. Adjusting the learning experience based on the detected stress level of the learner further adds to the potential for a learning experience that adapts to the learner, not the other way around.

5 Future Work

We have obtained ethical approval to conduct a comprehensive user study involving human participants. The study is designed to evaluate the impact of our biosignal-adaptive learning system compared to traditional non-adaptive XR learning approaches. The study will consist of two phases. In the first phase, participants engage in three English language learning modules of increasing difficulty within the XR environment. Throughout this phase, we collect physiological data (including heart rate, and potentially EEG and ECG in future iterations) alongside pre- and post-tests to assess language learning gains. Participants will also complete an *n*-back task to provide an additional measure of cognitive load.

The second phase leverages data gathered in the first phase to develop a machine learning model capable of classifying user stress levels in real-time. This model will then be integrated into the adaptive learning component, enabling the system to automatically adjust the difficulty of the learning modules based on the predicted stress level of the learner. By comparing results from both phases (including the learning outcomes, physiological responses, and subjective feedback of participants), we aim to quantify the effectiveness of our adaptive XR system in enhancing the language learning experience and improving learning gains compared to a traditional, non-adaptive XR-based approach.

6 Conclusion

This work contributes to the growing literature by presenting a system design for a biosignal-adaptive digital human tutor aimed at transforming language learning experiences. Our system integrates a customizable digital human within an immersive XR environment, driven by real-time physiological data from the learner. This system leverages the capabilities of LLMs and builds upon adaptive learning, embodied interaction, and affective computing. Our preliminary insights suggest that this approach holds potential for enhancing engagement and learning outcomes by dynamically adapting the educational content and feedback to the learner's cognitive and emotional state. The integration of biosignals into the interaction loop between the learner and the digital human tutor offers a pathway towards more personalized and effective learning experiences. This work underscores the transformative potential of biosignal-driven digital humans to reshape education, paving the way for a future where technology adapts to the learner in a truly human-centered manner.

References

[1] Liqaa Habeb Al-Obaydi. 2023. Humanistic learning elements in a blended learning environment: a study in an EFL teaching context. *Interactive Learning Environments* 31, 5 (2023), 3098–3111. https://doi.org/10.1080/10494820.2021.1919717

[2] Suad Alaofi and Seán Russell. 2022. The influence of foreign language classroom anxiety on academic performance in English-based CS1 courses. In *Proceedings of the 2022 Conference on United Kingdom & Ireland Computing Education Research*. ACM, New York, NY, United States, 1–7.

[3] Sara Ouald Chaib, Imane Joti, and Samira Khoulji. 2023. Evaluation of a Computer-Assisted Language Learning System Based on Adaptive Learning Designed for Self-training in Scientific French Language. *International Journal of Information and Education Technology* 13, 8 (2023).

[4] Luca Chittaro and Roberto Ranon. 2008. An adaptive 3D virtual environment for learning the X3D language. In *Proceedings of the 13th International Conference on Intelligent User Interfaces*. ACM, 419–420. doi:10.1145/1378773.1378846

[5] Jacobo Fernandez-Vargas, Tapio V. J. Tarvainen, Kahori Kita, and Wenwei Yu. 2017. Effects of Using Virtual Reality and Virtual Avatar on Hand Motion Reconstruction Accuracy and Brain Activity. *IEEE Access* 5 (2017), 23736–23750. doi:10.1109/access.2017.2766174

[6] Evelina Galaczi and Lynda Taylor. 2018. Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly* 15, 3 (2018), 219–236.

[7] Christina Gkonou. 2013. A diary study on the causes of English language classroom anxiety. *International Journal of English Studies* 13, 1 (2013), 51–68. https://doi.org/10.6018/ijes/2013/1/134681

[8] Nazlinur Gokturk and Evgeny Chukharev. 2024. Exploring the Potential of a Spoken Dialog System-Delivered Paired Discussion Task for Assessing Interactional Competence. *Language Assessment Quarterly* 21, 1 (2024), 60–99.

[9] David Griol, Ismael Baena, José Manuel Molina, and Araceli Sanchis de Miguel. 2014. *A Multimodal Conversational Agent for Personalized Language Learning*. Springer International Publishing, 13–21. doi:10.1007/978-3-319-07596-9_2

[10] Kunal Gupta, Yuewei Zhang, Tamil Selvan Gunasekaran, Nanditha Krishna, Yun Suen Pai, and Mark Billinghurst. 2024. CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Trans. on Visualization and Computer Graphics* 30 (2024), 2671–2681. Issue 5. doi:10.1109/TVCG.2024.3372130

[11] Elaine Hatfield, John T Cacioppo, and Richard L Rapson. 1993. Emotional contagion. *Current Directions in Psychological Science* 2, 3 (1993), 96–100.

[12] Ursula Hess and Agneta Fischer. 2014. Emotional mimicry: Why and when we mimic emotions. *Social and Personality Psychology Compass* 8 (2014), 45–57. Issue 2. doi:10.1111/spc3.12083

[13] Parneet Kaur, Harish Kumar, and Sakshi Kaushal. 2023. Technology-assisted language learning adaptive systems: a comprehensive review. *International Journal of Cognitive Computing in Engineering* 4 (2023), 301–313.

[14] Gholam Hassan Khajavy and Mahdieh Vaziri. 2024. Towards English language learners' wellbeing: Contributing factors and stressors. *ELT Journal* 78, 2 (2024), 117–126. https://doi.org/10.1093/elt/ccad048

[15] Shehroz S. Khan, Tong Zhu, Bing Ye, Alex Mihailidis, Andrea Iaboni, Kristine Newman, Angel He Wang, and Lori Schindel Martin. 2017. DAAD: A Framework for Detecting Agitation and Aggression in People Living with Dementia Using a Novel Multi-modal Sensor Network. In *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, New Orleans, LA, 703–710. doi:10.1109/ICDMW.2017.98

[16] Min-Soo Kim and Kwang-Seok Hong. 2016. An avatar expression method using biological signals of face-images. In *2016 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 1101–1103. doi:10.1109/ictc.2016.7763378

[17] Sueyoon Lee, Abdallah El Ali, Maarten Wijntjes, and Pablo Cesar. 2022. Understanding and Designing Avatar Biosignal Visualizations for Social Virtual Reality Entertainment. In *CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15. doi:10.1145/3491102.3517451

[18] Diane Litman, Helmer Strik, and Gad S Lim. 2018. Speech technologies and the assessment of second language speaking: Approaches, challenges, and opportunities. *Language Assessment Quarterly* 15, 3 (2018), 294–309.

[19] Gary J Ockey and Evgeny Chukharev-Hudilainen. 2021. Human versus computer partner in the paired oral discussion test. *Applied Linguistics* 42, 5 (2021), 924–944.

[20] Hong-Yi Pai, Ching-Huang Wang, and Yi-Chen Lai. 2024. Development of an Interactive Live Streaming System for Language Learning. *Journal of Internet Technology* 25, 1 (1 2024), 027–036. doi:10.53106/160792642024012501003

[21] K Rajitha and C Alamelu. 2020. A study of factors affecting and causing speaking anxiety. *Procedia Computer Science* 172 (2020), 1053–1058. https://doi.org/10.1016/j.procs.2020.05.154

[22] Claudia Repetto. 2014. The use of virtual reality for language investigation and learning. *Frontiers in Psychology* 5 (Nov 2014). doi:10.3389/fpsyg.2014.01280

[23] Sérgio Hortas Rodrigues, Samuel Mascarenhas, João Dias, and Ana Paiva. 2015. A process model of empathy for virtual agents. *Interacting with Computers* 27, 4 (2015), 371–391.

[24] Prasanth Sasikumar, Ryo Hajika, Kunal Gupta, Tamil Selvan Gunasekaran, Yun Suen Pai, Huidong Bai, Suranga Nanayakkara, and Mark Billinghurst. 2024. A User Study on Sharing Physiological Cues in VR Assembly Tasks. In *IEEE VR*. Institute of Electrical and Electronics Engineers Inc., 765–773. doi:10.1109/VR58804.2024.00096

[25] Tanja Schultz and Alexander Maedche. 2023. Biosignals meet Adaptive Systems. *SN Applied Sciences* 5, 234 (Aug 2023). doi:10.1007/s42452-023-05412-w

[26] Sfenrianto Sfenrianto, Yustinus B Hartarto, Habibullah Akbar, Mukhneri Mukhtar, Efriadi Efriadi, and Mochamad Wahyudi. 2018. An adaptive learning system based on knowledge level for English learning. *International Journal of Emerging Technologies in Learning (Online)* 13, 12 (2018), 191–200. https://doi.org/10.3991/ijet.v13i12.8004

[27] Veronika Timpe-Laughlin, Tetyana Sydorenko, and Phoebe Daurio. 2022. Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Computer Assisted Language Learning* 35, 5-6 (2022), 1194–1217.

[28] Veronika Timpe-Laughlin, Tetyana Sydorenko, and Judit Dombi. 2024. Human versus machine: Investigating L2 learner output in face-to-face versus fully automated role-plays. *Computer Assisted Language Learning* 37, 1-2 (2024), 149–178.

[29] Anita Vrins, Ethel Pruss, Jos Prinsen, Caterina Ceccato, and Maryam Alimardani. 2022. *Are You Paying Attention? The Effect of Embodied Interaction with an Adaptive Robot Tutor on User Engagement and Learning Performance*. Springer Nature, Cham, Switzerland, 135–145. doi:10.1007/978-3-031-24670-8_13