International Society of the Learning Sciences

LexiLearn: Sensory AI-enhanced VR Integration for Literacy Skills Acquisition

Marie P. P. M. Holm, Lucas Yeykelis, Anna C. M. Queiroz
mph718@miami.edu , lucasyeykelis@miami.edu , aqueiroz@miami.edu
University of Miami

Abstract: *LexiLearn* is an artificial intelligence (AI) enhanced virtual reality (VR) learning experience for literacy education. The system introduces a novel integration of AI with adaptable haptic interfaces, supporting both wearable gloves and handheld VR controllers. The project investigates how multisensory engagement within an adaptive, generative VR environment can support literacy acquisition, including spelling, reading through interactive and playful movement, phoneme–grapheme recognition and recall, and visual–sensory associations. Led by a cross-disciplinary research team, this project explores innovative alternatives to immersive learning environments. It examines how multisensory and AI-driven technologies can be meaningfully implemented for diverse learners, aiming to inform and transform the educational practice through practice-based design and research methods.

Introduction

Supporting students in the early stages of literacy acquisition remains an urgent challenge in global education. Literacy is a fundamental and critical prerequisite for equitable participation in society; ensuring access to effective literacy education is therefore essential for advancing human rights and enabling transformative life opportunities (UNESCO, 2019). However, current educational environments and digital interventions predominantly still rely heavily on visual and auditory modalities (Oakley et al., 2024). Yet, studies indicate that multisensory and embodied learning can enhance cognitive processing, memory, and engagement (Damsgaard et al., 2022; Fredens, 2018). Haptic feedback has been shown to convey information, support task performance, complement or substitute other sensory inputs, and contribute to a greater sense of immersion and realism in virtual environments (Culbertson et al., 2018). The integration of haptic devices has demonstrated positive effects on learning outcomes within STEM education (Orta Martinez et al., 2020). *LexiLearn* builds on these principles by investigating how a literacy-focused VR learning experience equipped with an adaptive and generative AI system, combined with wearable haptic gloves, potentially can support foundational skills such as letter recognition, spelling and phonics. The project is embedded within research practice partnerships involving schools, teachers, and students and is facilitated by a cross-disciplinary team of researchers from diverse fields, spanning design, psychology, education, and engineering. The remainder of this proposal outlines the project, research purpose, and future directions. With ethical approval secured from the Institutional Review Board (IRB), the study is now entering its co-design and participant user-testing phase. Initial insights and reflections on the design process will be presented at the ISLS 2026 conference.

Immersive virtual learning environments and play-based learning

Several studies and systematic literature reviews found that immersive virtual experiences enhanced learners' sense of presence, engagement, and conceptual understanding compared to traditional computer-based settings, with presence and self-efficacy mediating cognitive and affective learning outcomes (Chen, 2016; Makransky et al., 2020; Queiroz et al., 2023). These findings provide a conceptual basis for *LexiLearn*, investigating whether multisensory and adaptive features in immersive virtual experiences can further enhance literacy skills acquisition. Further studies found positive effects using VR experiences and haptics combined in procedural skill-related tasks that require the memorisation of steps or procedures, ranging from dental anaesthesia to piano lessons (Collaço et al., 2021; Pala & Mihci Turker, 2019). Given that spelling relies on procedural knowledge (Steffler, 2001), specifically the memorisation of phoneme–grapheme correspondences and letter sequences, the results from previous studies indicate promising potential for integrating multisensory and adaptive VR experiences into literacy learning domains.

The project employs elements of play such as narrative framing, embodied interaction, multisensory feedback, and reward cues to sustain motivation without fragmenting attention towards competitive metrics. The act of "harvesting words" in the orchard bridges playful exploration with literacy tasks, transforming the act of reading and spelling into a game-like VR experience. Furthermore, Fredens (2018) argues in his work on embodied learning that cognition emerges from the sensory–motor interplay between the brain, body, and environment. Emphasising that the hands stimulate the brain's frontal lobes, which are crucial for mental development and bodily experiences, form the conceptual foundations upon which linguistic development is built (Fredens, 2018). Drawing on these principles, Damsgaard et al. (2022) conducted a randomised controlled trial (the PLAYMORE Project) with children, demonstrating that whole-body movement paired with letter sound learning significantly improved children's letter sound knowledge (Damsgaard, 2022). These perspectives substantiate the exploration of immersive and embodied approaches within literacy learning. The shared experience of play also opens future possibilities for social learning, where future iterations of the project could allow collaborative participation in multiplayer settings, fostering co-learning and collective meaning-making based on principles of computer-supported collaborative learning (CSCL).

Literacy learning framework

The literacy learning framework proposed through *LexiLearn* is developed and designed on four core pillars: lexical development, creative visual association, phonetics and embodied learning and collaborative learning ecology, as described below:

Lexical development: Drawing upon cognitive science, we follow the literature that posits that the lexicon constitutes the primary locus for grammatical generalization across multiple representational tiers. Early, frequency-sensitive phonological pattern acquisition has shown to provide the essential mechanism for bootstrapping general grammatical competence in development (Beckman & Edwards, 2000). As learners engage with the words, visuals, phonemes and letters in *LexiLearn*, they practice decoding symbolic representations into sounds and meanings. The VR experience encourages the recognition of whole-word forms and repetition, supporting fluency and the transition from decoding to automaticity.

Creative Visual Association: The association between the written word and the images is based on the work in cognitive psychology and memory, specifically by demonstrating that linking a verbal stimulus (the word) with a non-verbal stimulus (the image) creates a more robust, multi-path memory trace (Paivio, 1990). *LexiLearn* offers visual cues and imagery to support recognition, memory, and comprehension by pairing each word with a corresponding AI-generated image of what the word represents. With this association, we expect the VR experience to support the development of the crucial orthographic-phonological-semantic mapping required for fluent reading (Warmington & Hulme, 2012).

Phonetics and embodied learning: Grounded in Embodied Cognition theory, body movement engagement has shown to facilitate encoding of lexical items by establishing non-abstract, action-based memory traces (Glenberg & Kaschak, 2002). This mechanism seems to be amplified by multisensory integration, wherein the temporal synchronisation of haptic, auditory, and visual input ensures the statistically optimal and highly reliable formation of complex perceptual-motor associations (Ernst & Banks, 2002). Building on this literature, *LexiLearn* associates visual, verbal and non-verbal stimuli, body movement, and the synchronisation between the vibrotactile feedback from the haptic gloves with phonetic audio on spelling skills.

Collaborative learning ecology: The efficacy of digital learning has been shown to be dependent on placing the educator in the centre of the experience's development, implementation and evaluation (Selwyn, 2021). The literature shows that data is only actionable when teachers can contextualise it to inform their professional judgment (Campen & Molenaar, 2020; Mandinach et al., 2006). Furthermore, there is a need for continuity between the digital and physical experiences, as effective learning requires a coherent curriculum that bridges and aligns technological activities with classroom goals, ensuring that embodied, hands-on learning is integrated rather than disrupted by digital tools (Mhlongo et al., 2023). In this context, we developed a dashboard that translates learners' activity in the VR experience into interpretable data, enabling educators to assess progress, identify individual strengths and challenges, and suggest learning tasks in the classroom accordingly. This aims to foster a reciprocal relationship between student and teacher, grounded in pedagogical reflection by evidence-informed insight.

Together, these pillars create a coherent, multisensory learning framework for exploring literacy acquisition through connecting the VR experience with activities in the classroom throughout the curriculum. The novelty of this research lies in its synthesis of multisensory, embodied, adaptive, collaborative and comprehensive dimensions, seeking to contribute to how multisensory digital experiences may potentially support learning pathways within literacy acquisition based on a model that is responsive to learners' cognitive, emotional, and sensory needs.

Design and development methods

The virtual reality experience and environment

LexiLearn is designed as an educational system composed of an adaptive and generative AI-enhanced VR learning experience coupled with sensory feedback facilitated as an educational game. The experience takes place in a calm, nature-inspired virtual orchard, designed to foster presence, curiosity, and sustained attention by providing a familiar, low-distraction context. Learners navigate freely through the space, selecting from a variety of fruit trees. When a tree is chosen, a target word appears alongside a generated image of the displayed word, along with fruits listed in front of the player containing letters as visual graphemes with associated phonetic delivery by audio cues linked to the interaction of picking up the “fruit-letters”. The learner engages in this embodied and interactive literacy task, grasping, aiming, and throwing the fruit-letters at a basket placed in front of the tree in the correct sequence to complete the word, where they have more letters to choose from than needed to complete the word (see Figure 1 for visuals of gameplay and link to video demo).

Visual and auditory feedback indicate if the letter selected is correct or not, and with a correctly completed word, a new word will appear with a new set of matching “fruit-letters”, fostering a continuous flow. We deliberately avoided scores and visible levels. This approach aims to cultivate a continuous, process-oriented flow, promoting learning through exploration instead of performance pressure. Aligning with the idea of the value of play lies not in goal attainment but in the process of exploration itself, which can support motivation, creativity, and engagement (Skovbjerg & Bekker, 2018). Before entering the orchard, learners encounter a customisation menu that allows them to tailor aspects of their experience. They can select thematic word categories (for example, animals, vehicles, emotions, etc.) to align with their interests, as well as choose background music, fostering resonant, personal, meaningful atmospheres. These were developed to foster agency, ownership, and intrinsic motivation, aligning with established research that more autonomous forms of motivation are associated with greater engagement, improved performance, and higher-quality learning (Ryan & Deci, 2000).

Figure 1
LexiLearn Visuals: Left, In-Experience View; Right, External View

Link to video demo with sound: https://youtu.be/VIuDrSX1wrA

Adaptive and generative AI-driven scaffolding

The system generates age-appropriate spelling words at calibrated difficulty levels (1–10), while DALL-E 3 produces corresponding visual imagery designed to reinforce meaning and word–image association of the displayed word (Li & Lan, 2022). Complementing these, a phoneme-level audio pipeline generates individual letter-sound pronunciations to support multi-sensory phonological learning. The system dynamically adapts by analysing each learner's accuracy and specific misspelling patterns at the character level. The first five words serve as a baseline assessment, prioritising age-appropriateness; thereafter, the system shifts to performance-driven calibration, increasing difficulty upon mastery (≥90% accuracy) and maintaining level when errors persist. By feeding prior misspellings directly into the generation prompt, the AI tailors subsequent word selection to target observed weaknesses — functioning as responsive scaffolding that mirrors a teacher's diagnostic adjustment (Harris et al., 2017). This adaptive mechanism aims to balance challenge and competence while accommodating individual differences in developing literacy. Through these parameters, the AI system functions as an intelligent mediator of literacy learning, guiding the sequencing of content and modulating task difficulty to align with each learner’s evolving proficiency.

Sensory integration

The haptic feedback is delivered through gloves equipped with linear resonant actuators (LRAs) that provide localised vibrotactile stimulation to the hand and fingertips. Hand and finger movements are captured via inertial measurement unit (IMU) sensors, in conjunction with optical hand tracking from the VR headset (Meta Quest 3). This multimodal integration facilitates context-responsive vibrotactile feedback aligned with literacy-oriented tasks within the VR experience. Rather than replicating physical tactile sensations in a one-to-one correspondence with the phonetic audio, the haptic system is designed to utilise more systematic and differential haptic cues. We found, during the design process, that audio-based haptic pairing with phonetic sounds was too subtle for the users to justifiably notice. Through a systematic integration of vibrotactile feedback, this project seeks to support grapheme–phoneme associations using clearly differentiable and intuitive sensory stimuli, while investigating the impact of haptic feedback on learning outcomes.

Design and curriculum integration

LexiLearn aims not only to create a single standing VR learning experience but also to explore a holistic, comprehensive learning experience that connects the embodied VR learning experience with in-class curriculum activities supporting literacy skills. In this context, we developed a dashboard combined with the *LexiLearn* VR experience that can be accessed through computers, tablets or mobile phones. The dashboard provides educators with access to data recorded from each student interaction within the VR experience. They can see how long and how often the learners played, the letters practised and the learners' accuracy and mistakes. The AI provides insights into the mistakes' patterns (for example, if the learner frequently selected "b" instead of "d") and suggests personalised student activities that the teacher can implement in the classroom to support the specific skill level of each learner. This way, the VR learning experience is not isolated from the curriculum but integrated with classroom activities. The dashboard supports a continuum of personalised learning throughout the year and functions as a bridge between learner experience and teaching practice. Insights from immersive sessions are translated into actionable pedagogical reflection, aligning with broader efforts to support evidence-informed teaching. In this context, we aim for the AI system to be positioned as a complementary pedagogical tool rather than a substitute for classroom teaching, extending formative assessment practices and supporting personalisation within diverse and mixed-ability classrooms.

Conclusion and future directions

LexiLearn is a comprehensive system that integrates a novel, adaptive AI-enhanced VR learning experience with vibrotactile haptic feedback and a pedagogical data dashboard. The synthesis of adaptive AI-driven content sequencing, synchronised grapheme-phoneme-tactile cueing, and data translation mechanism provides a robust framework for investigating orthographic-phonological-semantic mapping and comprehensively embedding the digital learning experience in the classroom. Future research must prioritise the empirical validation of this integrated system, investigating the differential effects of the multisensory condition on measures of spelling proficiency, phonological awareness, and lexical development. Specifically, longitudinal analyses should quantify the relationship between the dashboard's data analytics—including machine-identified mistake patterns—and the subsequent impact of educator-implemented, evidence-informed instructional modifications on the sustained continuity and alignment across the digital VR learning experience and the curriculum.

Interactive demo session

The researchers will provide attendees with a live, hands-on demonstration of the *LexiLearn* learning system, including the teacher's dashboard. Attendees will be able to experience the VR learning experience directly, exploring how vibrotactile feedback provided by the haptic gloves is coupled with visual and auditory cues to support embodied interaction and literacy learning. Additionally, they will be invited to contribute their feedback on the experience and the dashboard to support further potential applications and broader reach across the educational practice. The aim for the demo session is to engage attendees in trialling the novel set-up and spark curiosity while supporting ISLS's knowledge-sharing space within educational research. The session will also allow the research team to share formative findings from early user-testing with teachers and students. Ultimately, this demo aims to foster dialogue around the integration of new immersive technologies in education, encouraging collaboration, reflection, and new directions for future research and educational pathways.

References

Beckman, M. E., & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child development, 71*(1), 240-249.

Campen, K. V., & Molenaar, I. (2020). How Teachers Integrate Dashboards into Their Feedback Practices. *Frontline Learning Research, 8*(4), 37-51. https://doi.org/10.14786/flr.v8i4.641

Chen, YL. The Effects of Virtual Reality Learning Environment on Student Cognitive and Linguistic Development. *Asia-Pacific Edu Res* 25, 637–646 (2016). https://doi.org/10.1007/s40299-016-0293-2

Collaço, E., Kira, E., Sallaberry, L. H., Queiroz, A. C. M., Machado, M. A. A., Crivello Jr., O., & Tori, R. (2021). Immersion and haptic feedback impacts on dental anesthesia technical skills virtual reality training. *Journal of Dental Education, 85*(4), 589-598. https://doi.org/10.1002/jdd.12503

Corbalan, Gemma & Kester, Liesbeth & Van Merrienboer, Jeroen J. G.. (2008). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. Contemporary Educational

Psychology. 33. 733-756. https://doi.org/10.1016/j.cedpsych.2008.02.003
Culbertson, H., Schorr, S. B., & Okamura, A. M. (2018). Haptics: The present and future of artificial touch sensation. *Annual Review of Control, Robotics, and Autonomous Systems, 1*(1), 385–409.
Damsgaard, L., Nielsen, A.-M. V., Gejl, A. K., Malling, A. S. B., Jensen, S. K., & Wienecke, J. (2022). Effects of 8 weeks with embodied learning on 5-6-year-old Danish children's pre-reading skills and word reading skills: The PLAYMORE Project, DK. *Educational Psychology Review*, 34(3), 1709-1737.
Damsgaard, L., Eiberg, M., Jensen, C. D., & Melby, P. S. (2022). Whole-body movement and letter–sound knowledge in early literacy acquisition: Results from the PLAYMORE Project. *Journal of Educational Psychology, 114*(3), 459–473. https://doi.org/10.1037/edu0000625
Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(6870), 429-433. https://doi.org/10.1038/415429a
Fredens, K. (2018). *Læring med kroppen forrest: En hjernebrugsbog*. Akademisk Forlag.
Gallagher, L., Means, B., & Padilla, C. (2008). *Teachers' use of student data systems to improve instruction: 2005 to 2007* (NCEE 2009-4075). U.S. Department of Education, Office of Planning, Evaluation and Policy Development.
Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. Psychonomic bulletin & review, 9(3), 558-565.
Harris, K. R., Graham, S., Aitken, A. A., Barkel, A., Houston, J., & Ray, A. (2017). Teaching spelling, writing, and reading for writing: Powerful evidence-based practices. *Teaching Exceptional Children*, *49*(4), 262-272. https://doi.org/10.1177/004005991769725
Holmes, Wayne & Porayska-Pomsta, Kaska & Holstein, Ken & Sutherland, Emma & Baker, Toby & Shum, Simon & Santos, Olga C. & Rodrigo, Mercedes & Cukurova, Mutlu & Bittencourt, Ig & Koedinger, Kenneth. (2021). Ethics of AI in Education: Towards a Community-Wide Framework. International Journal of Artificial Intelligence in Education. 32. https://doi.org/10.1007/s40593-021-00239-1
Li, P., & Lan, Y. J. (2022). Digital language learning (DLL): Insights from behaviour, cognition, and the brain. *Bilingualism: Language and Cognition, 25*(3), 361-378. doi:10.1017/S1366728921000353
Mandinach, E. B., Honey, M., & Light, D. (2006, April). A theoretical framework for data-driven decision making. *In annual meeting of the American Educational Research Association*, San Francisco, CA (pp. 39-52).
Makransky, G., Petersen, G.B. and Klingenberg, S. (2020), Can an immersive virtual reality simulation increase students' interest and career aspirations in science?. Br. J. Educ. Technol., 51: 2079-2097.
Mhlongo, S., Mbatha, K., Ramatsetse, B., & Dlamini, R. (2023). Challenges, opportunities, and prospects of adopting and using smart digital technologies in learning environments: An iterative review. *Heliyon, 9*(6). https://doi.org/10.1016/j.heliyon.2023.e16348
Oakley, G. (2024). A scoping review of research on the use of digital technologies for teaching reading fluency. *Education Sciences*, *14*(6), 633. https://doi.org/10.3390/educsci14060633
Orta Martinez, C. M. Nunez, T. Liao, T. K. Morimoto and A. M. Okamura, "Evolution and Analysis of Hapkit: An Open-Source Haptic Device for Educational Applications," in *IEEE Transactions on Haptics*, vol. 13, no. 2, pp. 354-367, 1 April-June 2020, https://doi.org/10.1109/TOH.2019.2948609
Pala, F. K., & Mihci Turker, P. (2019). Developing a Haptic Glove for Basic Piano Education. *World Journal on Educational Technology: Current Issues*, *11*(1), 38-47.
Paivio, A. (1990). Mental representations: A dual coding approach. Oxford university press.
Queiroz, A. C. M. (2023). *Collaborative tasks in immersive virtual reality increase …* (ISLS Annual Meeting 2023, pp. 27-34). International Society of the Learning Sciences.
Que, Ying & Cheng, Lang & Zhang, Tianshu & Chan, Carol & Hu, Xiao. (2025). Comparing Collaboration Patterns in Virtual Reality Content Co-Creation Activities between High- and Low-Performing Secondary School Students. https://doi.org/10.22318/cscl2025.742933
Ryan, Richard & Deci, Edward. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. American Psychologist. 55. 68-78.
Selwyn, N. (2021). *Education and technology: Key issues and debates*. Bloomsbury Publishing.
Skovbjerg, Helle Marie & Bekker, Tilde. (2018). *The value of play*. Design School Kolding.
Steffler, D. J. (2001). Implicit cognition and spelling development. *Developmental review*, *21*(2), 168-204.
UNESCO. (2019). *UNESCO strategy for youth and adult literacy (2020–2025)* (40 C/25). United Nations Educational, Scientific and Cultural Organization.
Warmington, M., & Hulme, C. (2012). Phoneme awareness, visual-verbal paired-associate learning, and rapid automatized naming as predictors of individual differences in reading ability. *Scientific Studies of Reading, 16*(1), 45-62. https://doi.org/10.1080/10888438.2010.534832

