

Emotion-Aware Artificial Intelligence Systems for Enhancing English Learning Engagement

Alaviyya Bahruz Nuri^{1*}

^{1*}PhD Student, English Teacher, Nakhchivan State University, Nakhchivan, Azerbaijan.

E-mail: eleviyyenuri@ndu.edu.az, Orcid: <https://orcid.org/0000-0003-4219-1071>

Abstract: The study is driven by the observation that many computer-assisted language learning (CALL) resources lack consideration of the affective elements of language learning and that many learners are regarded as rational information processors. This paper explores the theoretical principles, design principles, and pedagogic promise of emotion-aware artificial intelligence (AI) systems to recognize, understand, and react to the emotions of learners in real-time to maintain their engagement in English as a Second/Foreign language (ESL/EFL) learning. The paper draws from the literature on affective computing, control-value theory and second language acquisition studies to integrate the latest findings from the field with existing review studies on multimodal emotion recognition, AI-based emotional intelligence, and language learning with chatbots. A conceptual emotion-aware artificial intelligence framework is presented, consisting of the multimodal sensing layer (facial expression, speech prosody, text or keystroke behavior), the emotion recognition module, the learner emotion and learner engagement model, and the adaptive pedagogical response engine that dynamically adapts the difficulty, pacing, and feedback tone of a task based on the inferred affective state. The paper also addresses how foreign language anxiety can be addressed in such systems, how they can keep learners in the optimal zone of enjoyment and flow, and how they can scaffold for learners with varying emotional profiles. The reviewed literature suggests that emotion-informed AI interventions can lead to a decrease in speaking anxiety and an increase in self-reported engagement and learning outcomes compared to emotion-uninformed interventions, with effect sizes differing depending on the specific type and mix of emotional and cognitive support provided. Finally, the paper outlines several challenges for the implementation, such as issues of data privacy and differences in emotion expression across cultures, algorithmic bias, and the potential for relying too heavily on automated emotion detection, and provides directions for future empirical research on emotion-aware AI in real English language learning settings.

Key Words: emotion-aware artificial intelligence, affective computing, english language learning, learner engagement, multimodal emotion recognition, adaptive feedback, foreign language anxiety

(Received: 24 March 2025; Revised: 14 April 2025; Accepted: 02 June 2025; Published: 30 June 2025)

Introduction

English has been, and will remain, the global language of academic, professional and digital communication, and the need for effective English as a Second/Foreign Language (ESL/EFL) instruction is increasing in formal and informal educational settings. In recent years, Artificial Intelligence (AI) has increasingly found its way into the field of English language teaching, as in the form of intelligent tutoring systems, conversational agents, automatic system of writing assessment, and adaptive content platforms. The systems have tended to be designed so as to model the states of the learner's cognitive system rather than the learner herself, and to view the individual as a machine-like processor of linguistic input without emotions. But significant second language acquisition (SLA) research presents empirical evidence of the

systematic effect of emotional responses including enjoyment, anxiety, boredom and frustration on attention, willingness to communicate and final learning outcomes.

The technical basis for machines that can recognize and respond to the human affect was laid by affective computing, defined as computing that relates to, arises from, or deliberately influences emotion (Picard, 2003). With the progress in computer vision, speech processing and physiological sensing, emotion-aware systems have since been applied in general educational technology to estimate learner engagement from facial expressions, to augment the estimation of learner engagement using facial and physiological signals, and to estimate learner engagement based on behavioral interaction patterns (Whitehill et al., 2014; Monkaresi et al., 2016; Altuwairqi et al., 2021). The use of these techniques, and more generally AI-based emotional intelligence, has started in English language education to alleviate anxiety associated with spoken language, to maintain attention, and to offer tailored feedback (Chandel & Lim, 2024; Liu et al., 2024; Wang et al., 2024).

While these advances have been made, there is still a fragmented literature on emotion recognition, with its development largely occurring within the general educational technology and human–computer interaction community, and other studies have concentrated on discrete tools such as chatbots or generative AI writing assistants, but not within an integrated architecture that seamlessly connects real-time affect detection with pedagogical adaptation. In this paper, it aims to fill this gap by (a) reviewing and summarizing existing literature on emotion, engagement, and AI in the field of English language education; (b) outlining an evaluation methodology that is consistent with previous empirical designs; and (c) discussing the pedagogical implications, challenges, and ethical considerations of implementing such systems in real classrooms.

The rest of this paper is organized as follows. This paper reviews literature on emotion, engagement and L2 acquisition, emotion and affective computing, emotion and multimodal emotion recognition and the emerging role of AI in EI in L2 learning in order to highlight the specific gap in the literature that is the subject of this paper. Section 3 introduces the proposed emotion-aware AI framework, which outlines the design principles, overall system architecture, and the multimodal emotion detection module, the learner emotion and engagement model, the adaptive pedagogical response engine, and the longitudinal feedback loop. Section 4 describes a proposed mixed methods evaluation methodology for testing the framework in authentic classrooms. The elements of the framework are evaluated based on the gathered literature in Section 5, with the expected effectiveness and boundary conditions being assessed. Section 6 provides pedagogical and practical implications for curriculum designers, teachers, and institutions; and Section 7 discusses issues of implementation and ethical considerations such as data privacy, Crosscultural variability, and the risk of over-reliance on automated affect detection. The paper is concluded in section 8, which gives directions for further research.

Literature Review

Emotion, Engagement, and Second Language Acquisition

Engagement in language learning is usually viewed as a multi-dimensional construct that includes behavioral, cognitive and emotional dimensions, with the latter (interest, enjoyment, sense of belonging) as a prerequisite for the former. In SLA research, the positive psychology movement has turned research's focus to foreign language enjoyment (FLE), a construct that is different from, and is not the opposite of, foreign language classroom anxiety (FLCA) (Dewaele & MacIntyre, 2014). Empirical research has demonstrated that enjoyment and anxiety are compatible (Dewaele & Alfawzan, 2018; Dewaele et al.,

2024), that both are affected by the teacher's behavior and classroom climate (Dewaele & Alfawzan, 2018; Fryer et al., 2017) and that enjoyment can even outweigh the interference of anxiety in performance when the context is right (Fryer et al., 2019). Later research was conducted to determine how enjoyment, anxiety and boredom changed measurably when transitioning between online and offline modes of instruction in the emergency remote teaching context (Dewaele et al., 2024). The overall results indicate that when developing an AI-based tool to facilitate the learning engagement in English, it is not only necessary to pay attention to negative affect (anxiety, boredom, frustration) but also to foster positive affect (enjoyment, interest, flow).

Affective Computing and Multimodal Emotion Recognition

Affective computing scholarship has outlined the fundamental technical challenges regarding the recognition, expression and response to affect in systems such as ambiguity of emotion expression, individual and culture differences, and the ability to acquire reliably labeled training data (Picard, 2003). In the educational domain, in particular, classifiers based on facial expressions have been developed to detect discrete academic emotions and levels of engagement from video streams with some accuracy, approaching that of trained human coders (Whitehill et al., 2014). Further studies have added physiological data, such as heart rate, to the task of inferring engagement, using facial expressions alongside the physiological signal to make the engagement detection more robust (Monkaresi et al., 2016); other work has used behavioral and interaction-log data (including mouse movements, response times, and task-switching) to provide an indirect measurement of engagement without the need for ongoing video recording (Altuwairqi et al., 2021). It was found in early multimodal work that probabilistic sensor-fusion models of combining two or more modalities (facial expression, posture, and on-task computer activity) could classify learner interest more accurately than any individual channel alone. To complement these detection-oriented investigations, a theoretical model of affect dynamics in complex learning was proposed which posited that when cognitive impasses arose there would be some students who would move into states of confusion and others who would move into states of engagement/flow and that resolving confusion, not avoiding it, was often associated with deeper learning (Blikstein & Worsley, 2016). All together, this collection of work lays the groundwork for the technical feasibility of automated emotion and engagement detection and a theoretical framework on how affective states change in the learning process that are pertinent to the design of an emotion-aware AI for English instruction.

AI-Driven Emotional Intelligence in English Language Education

In recent years, an increasing number of systematic reviews have focused directly on the interaction between emotion, AI, and English language learning and teaching. The empirical research of the AI-assisted ESL/EFL learning reviewed indicated that there were two emotions that were most often investigated in the literature: enjoyment and anxiety; and that AI tools could influence emotions more so for task design through personalization, adaptability, and interactivity than by providing generic tech exposure (Chandel & Lim, 2024). The first three of these five affordances (human-like conversation, personalized real-time feedback, translation of visual input into English text, generation of personalized content, and recognition or analysis of learner emotion) involve a combination of cognitive and emotional support and they were linked to more consistent improvements in behavioral, cognitive and affective outcomes, whereas the last affordance (emotional support) was not (Liu et al., 2024). A systematic review of the literature on AI tools for emotion assessment in L2 learning environments in general showed that there is potential for using AI tools to assist socio-emotional learning, but they need to be further validated in various learning environments (Rojas Vistorte et al., 2024). As a specific example of teacher preparation, there have already been some first steps towards the transformation of classroom practice with the use of

AI tools such as chatbots and pronunciation feedback applications in L2 foreign language learning, but their use in teacher preparation is in its early stages with equity and evaluation still to be developed (Peña-Acuña & Corga Fernandes Durão, 2024).

This literature has been enhanced by another layer of generative AI. A qualitative study on the use of generative AI in L2 writing found that these tools can support the emotional aspects of teaching, such as reducing students' fear of writing in an L2 context because they feel like they are in a safe, non-judgmental environment. (Wang et al., 2024) The systematic review on the uses of ChatGPT in ESL/EFL also reported a rising number of studies that demonstrated positive affective outcomes, such as boosting student confidence and alleviating anxiety, and highlighted the need for further systematic and empirical research with longer time frames (Lo et al., 2024). In support of this generative AI literature, previous experimental studies on chatbot task partners in language courses showed that, while chatbot novelty can boost learner interest in the short term, the interest of the language course itself can be sustained by the chatbot's perceived competence along with the design of the conversational task, rather than by the novelty of the chatbot itself (Fryer et al., 2017; Fryer et al., 2019).

This literature leaves three gaps. First, emotion-recognition methods that have been successfully implemented in general pedagogical technology (facial expression, physiological and behavioural-log based recognition) are seldom used in a unified architecture specifically designed for English language skill areas like speaking, listening, collaborative writing and where affective states like speaking anxiety play a particularly important role. Secondly, most studies that investigate an ELS only test one tool (chatbot, a generative AI writing assistant), and not an entire closed-loop system, where detected affect continuously and directly influences pedagogical decisions like task difficulty, pacing, and feedback tone. Third, previous reviews have repeatedly emphasized the need for longitudinal, classroom-embedded assessment, but empirical investigations of the validation of integrated EAI systems for assessing English learning engagement are limited. The framework presented in Section 3 aims at rectifying these gaps by providing a multimodal detection system based on the validated methods and an adaptive pedagogical engine for the types of tasks in English.

Proposed Emotion-Aware AI Framework for English Learning

Design Principles

The proposed framework is based on four design principles, drawn from the literature reviewed. First, multimodality: There is no single affective channel which is infallible, and the system should make use of facial expression, speech-prosodic and text or keystroke behaviors as a combination, because multimodal fusion is better than unimodal fusion in terms of detection accuracy (Monkaresi et al., 2016). Second, pedagogical actionability: emotion detection is not considered as an end in itself, but rather as a step that needs to lead to a tangible pedagogical adjustment, which is the same as what was found when emotion detection was added to cognitive support, resulting in better outcomes than emotion detection alone (Liu et al., 2024). Third, affect-dynamics awareness: affect is not treated as if it must be eradicated as it comes up, but rather aware of the difference between productive confusion (which is a precursor to deeper learning) and unproductive frustration or boredom (which is a precursor to disengagement), based on the affect-dynamics model of complex learning (Blikstein & Worsley, 2016). Fourth, learner-centred transparency: with well-founded concerns about the collection of biometric data in education, the framework introduces explicit consent and transparency mechanisms which inform learners about what is being sensed and for what purpose.

System Architecture

The proposed system architecture is illustrated in figure 1, with five elements: (1) multimodal sensing layer, which passively captures the learner's facial video, microphone audio, and text/keystroke interactions during English learning activities; (2) emotion recognition module, which classifies the captured signals into the discrete or dimensional affective states; (3) learner emotion and engagement model, which builds an engagement index as a running score and an emotion profile as a long-term sequence of affect classifications; (4) adaptive pedagogical response engine, which maps the engagement and affective state onto specific pedagogical adjustments; and (5) a longitudinal learner-profile and analytics store, which enables real-time adaptation and post-hoc evaluation by teachers and researchers.

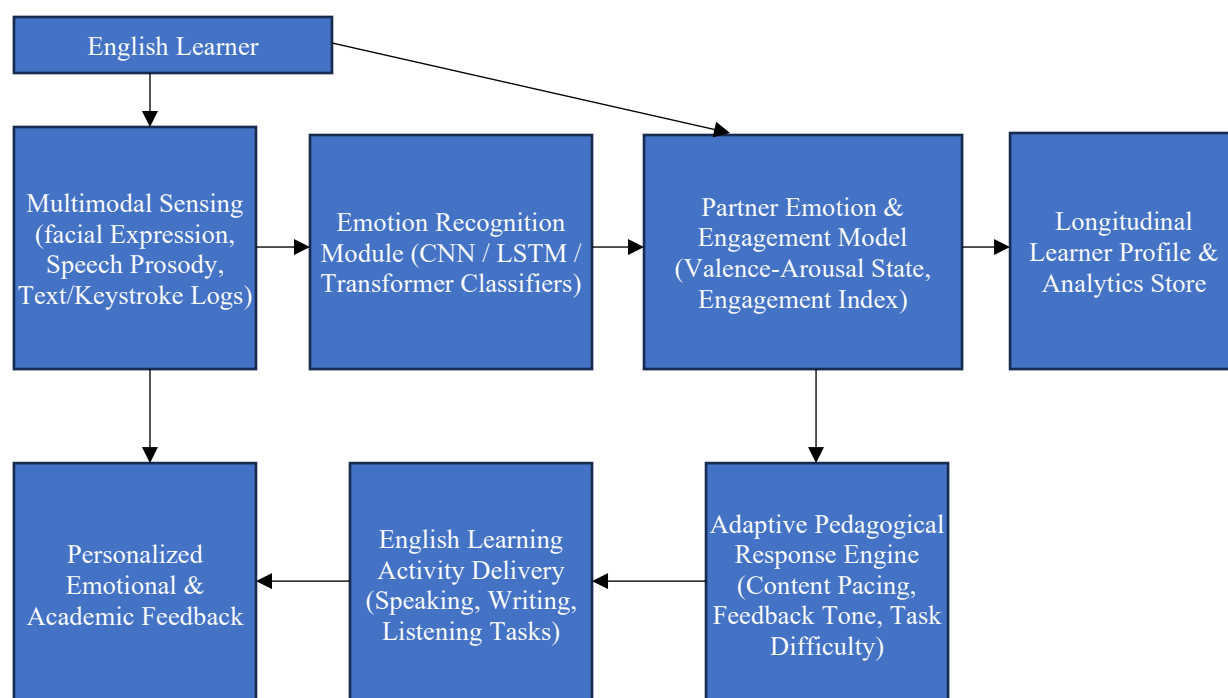


Figure 1: Proposed Architecture of an Emotion-Aware AI System for English Learning Engagement.

Multimodal Emotion Detection Module

The detection module also contains a facial-expression-based engagement classifier, similar to those used in previous research which are based on a convolutional network trained to detect academic emotions that are relevant to English learning, including confusion while listening to a teacher or peers, frustration while speaking to a teacher or peers, and enjoyment when working collaboratively (Whitehill et al., 2014). The module relies on the behavioral-engagement-detection method which has been proven effective in online learning environments for recognizing webcam users, as well as for recognizing speech-prosodic and text-production features of learners' utterances during speaking tasks and writing tasks, respectively (Altuwairqi et al., 2021). Combining them together in a single momentary effect estimate, as shown in the late-fusion classifier on the sensor-fusion approach used for learner interest detection and on the approach for fusion of facial and physiological features for engagement detection is a late-fusion approach (Monkaresi et al., 2016).

Learner Emotion and Engagement Model

Momentary affect estimates are condensed across a sliding time window into (a) a continuous engagement index, and (b) a categorical emotional state (e.g., flow/engaged, confused, anxious, bored,

frustrated), similar to affect-transition patterns noted in complex learning (Blikstein & Worsley, 2016). Following the evidence that foreign language enjoyment and anxiety can be triggered by specific classroom variables and by specific task types and by specific interaction partners, the model keeps a longitudinal learner profile of recurrent triggers of anxiety or disengagement, such as specific task types (spontaneous speaking) or interaction partners (synchronous peer feedback) (Dewaele & MacIntyre, 2014; Dewaele & Alfawzan, 2018).

In order to act out this categorical state, the model assigns momentary affect estimates to a two-dimensional valence–arousal space (illustrated conceptually in figure 2). Task-relevant emotional experiences (e.g., speaking anxiety, listening confusion) are in the high-arousal, negative-valence quadrant; collaborative task enjoyment and flow are in the high-arousal, positive-valence quadrant; post-task satisfaction is in the low-arousal, positive-valence quadrant; and disengaged boredom during repetitive drills is in the low-arousal, negative-valence quadrant. This placement would lead to difficulty reduction or switch to a lower-stakes interaction partner for responses in the high-arousal, negative-valence quadrant and increase task novelty or pacing for responses in the low-arousal, negative-valence quadrant (as per affect-dynamism theory, which states that difficulty reduction or shift to a lower stakes interaction partner is warranted for responses in low-arousal, negative-valence quadrant, whilst increased task novelty or pacing is warranted for responses in high-arousal, negative-valence quadrant) (Blikstein & Worsley, 2016).

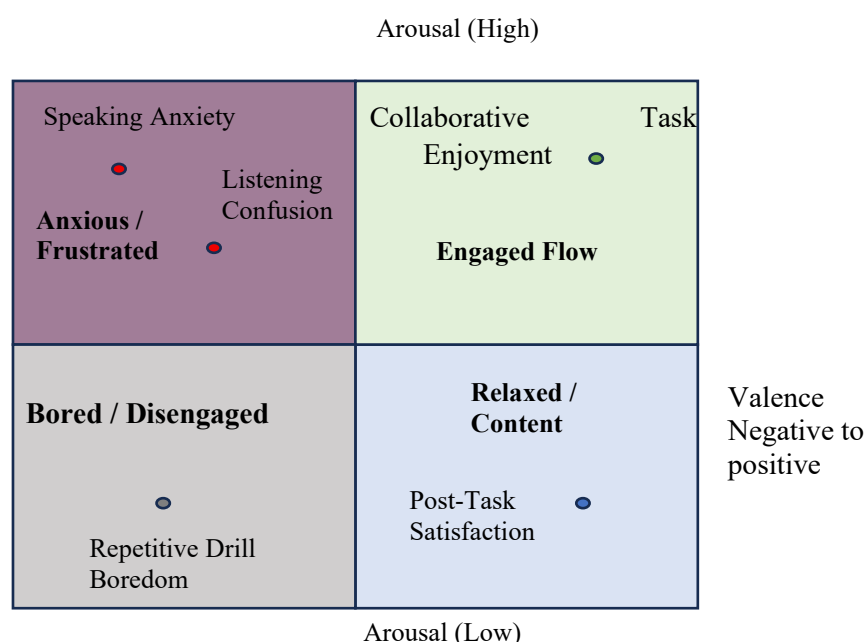


Figure 2: Conceptual Valence–Arousal Placement of Learner Affective States Used by the Engagement Model, with Illustrative English-Learning Examples.

Adaptive Pedagogical Response Engine

The five possible instructional actions are: Difficulty Pacing Change, Interchange Partner Change, Feedback Tone Change, and Affective Check-in. This design directly applies the principle that cognitive adaptation (task/content adaptation) combined with emotional support (feedback tone, encouragement) will have more consistent gains than emotional support alone (Liu et al., 2024). The engine uses evidence that sustained interest in language tasks that are mediated by chatbots is related to perceived competence, not just novelty (Fryer et al., 2017; Fryer et al., 2017), and so when it is detected that the user is disengaged

from a task during chatbot use, it focuses on making the task more competent than it is affectively, not simply because of its novelty (Fryer et al., 2019).

Feedback and Longitudinal Analytics Loop

In addition to immediate adaptation, the architecture affects and engagement trajectories are also stored in a longitudinal analytics store that is accessible to teachers, which means that any patterns that emerge over weeks (e.g., a teacher may notice that a student's speaking anxiety is persisting from week to week) that can be detected outside of the automated system can trigger a human response. This loop is based on the recommendations of systematic reviews of emotion assessment using AI which suggest that emotion assessment tools should supplement rather than supplant humans' judgement regarding learners' socio-emotional development (Rojas Vistorte et al., 2024).

Proposed Evaluation Methodology

A quasi-experimental, mixed-methods approach should be used to evaluate the proposed framework as in previous empirical investigations in this area. The experimental group employing the emotion-aware AI system would be contrasted with a control group that employs the equivalent emotion-agnostic version of the same system, consistent with the designs employed to isolate the effect of using AI to teach emotional intelligence in speaking. All quantitative measures should include: (a) validated FLE and FLCA scale scores pre- and post-intervention (Dewaele & MacIntyre, 2014; Dewaele & Alfawzan, 2018); (b) an automatically logged engagement index derived from the system itself; (c) task-level learning outcomes, if applicable, appropriate to the skill being trained (e.g., speaking fluency ratings, writing quality scores); and (d) classifier accuracy against human-annotated affect labels, as done with facial-expression based engagement detectors (Whitehill et al., 2014), using the approach to validation. Qualitative included in the study would involve learner interviews or open-ended surveys asking about how helpful and comfortable students felt with the biometric sensing, and how authentic they felt AI was in its emotional responses, as happened in previous studies of generative AI in the field of language learning (Wang et al., 2024). The minimum time period for evaluation is recommended to be one academic term; the significance of FLE and FLCA being not constant after termination of one session has also been demonstrated across a course (Dewaele et al., 2024).

5. Synthesis of Evidence and Discussion

Table 1 summarises the key outputs of the literature review and how this supports design and anticipated impact of the proposed framework.

The three patterns that emerge in the synthesized evidence are. First, the technologies of detecting emotion and engagement that have been validated in other fields than ELT (e.g., facial-expression classifiers, physiological-facial fusion, behavioural-log based detection (Altuwairqi et al., 2021) or multimodal sensor fusion) are technically mature enough to be adapted for use in ESL/EFL settings with the only requirement being retraining them on English-learning behaviours such as hesitation in speaking and confusion in listening. Second, the literature in the English-language domain is converging towards the conclusion that affective support is most effective when combined with cognitive or pedagogical support, rather than with simply affective support, which directly motivates the closed-loop design in Section 3, where detected affect is used to provide concrete support for instruction rather than generic encouragement. Third, there is a clear engagement effect of novel AI tools like chatbots and generative AI assistants that seem to be temporary unless coupled with actual improvements in perceived competence,

highlighting the need for longitudinal analytics loop to differentiate between short-term engagement and longer-lasting gains in perceived competence.

Table 1: Representative Studies Informing the Proposed Emotion-Aware AI Framework

Study	System / AI Technique	Target Skill / Context	Key Affective / Engagement Finding
Liu et al. (2024)	Systematic review of emotional AI applications	EFL classrooms, conversational agents	Five affordances identified; emotional support combined with cognitive support outperforms emotional support alone
Vistorte et al. (2024)	Systematic literature review of AI emotion-assessment tools	General learning environments	AI-based emotion assessment supports socio-emotional development but requires validated, context-sensitive models
Whitehill et al. (2014)	Facial-expression classifiers (engagement labels)	Online video-based learning	Automatic recognition of student engagement from facial expressions achieves accuracy comparable to human observers
Monkaresi et al. (2016)	Facial-expression + heart-rate fusion	Structured writing activity	Combining physiological and facial signals improves engagement-detection accuracy over single-modality models
Fryer et al. (2017) (2019)	Conversational chatbot task partners	EFL conversation practice	Chatbot novelty briefly raises interest, but sustained interest depends on perceived competence and task design, not affect alone

These patterns also outline the boundary conditions of the effectiveness of the framework. Since foreign language enjoyment and anxiety are influenced by factors of the teacher and the classroom factors that an AI system cannot completely regulate, the framework is regarded more as a support to, rather than a replacement for, human instructional relationships (Dewaele & MacIntyre, 2014; Dewaele & Alfawzan, 2018; Dewaele et al., 2024). Likewise, as the classifiers need to be further validated in different educational and cultural contexts, deployment should be done carefully and the classifiers' performance should be monitored by learner subgroups before the fully automated decisions of instructional action based on the emotion assessment to be taken (Rojas Vistorte et al., 2024).

Pedagogical and Practical Implications

From the perspective of curriculum designers, the framework suggests that English language content should be presented in a type of courseware that contains clearly defined affective branch points (alternate and varying versions of the same task, each with a different amount of difficulty and interactional style) for the adaptive engine to choose from, not just a single task sequence. The longitudinal analytics loop provides teachers with a diagnostic tool for recognizing patterns of speaking anxiety and disengagement among learners that are unlikely to be successfully addressed by automated adaptation, and offer opportunities for human intervention at points where it is likely to be most effective. A recent study

demonstrates that in the context of cognitive-emotional support, it is pedagogical adaptivity, rather than affect-recognition accuracy, that is the most important factor for institutions to consider when planning to implement an AI language-learning tool in their classrooms.

Challenges and Ethical Considerations

The implementation of emotion-aware AI in an English language classroom poses multiple challenges that need to be solved before the technology can become commonplace. One of the key concerns is data privacy, as the logs of facial videos, speech, and behavior represent sensitive biometric data that needs to be collected with informed consent and in accordance with relevant data protection laws and regulations. There is also a challenge of cross-cultural and individual variability in emotion expression, with facial and prosodic emotion cues that are validated across one population not necessarily being generalizable across different linguistic and cultural populations without the risk of systematic misclassification and algorithmic bias. It is also possible that learners might think that continuous affective sensing is being used for monitoring, rather than for support, and in this way, it could create anxiety and jeopardize the engagement it is supposed to promote. Last, but not least, using automated affect detection can risk de-skilling the skill of teaching to read the emotional climate of the classroom, which is why the framework considers it a decision-support input for teachers and not an authority on learners' emotion. Solving these challenges will necessitate transparent design, opt-in sensing, frequent bias audits of classifiers for various subgroups and continued teacher engagement with system output and action.

Conclusion

This paper has proposed that the emotional status of learners should be considered not just a by-product but as a major part of the English language learning design, requiring AI systems to consider emotional state as their key focus. The paper was based on the foundations of affective computing (Picard, 2003) and emotion and engagement detection methods that have been proven in multimodal emotion detection methods and validated in multimodal emotion detection methods in the field of SLA in the area of foreign language enjoyment and anxiety detection, and the findings of the systematic review that were conducted in the field of emotional AI and generative AI in ELL. The paper is based on the foundation of affective computing (Picard, 2003) and validated emotion and engagement detection techniques and findings of the systematic reviews conducted in the field of emotional AI and generative AI in ELL. The synthesized evidence suggests that, overall, an emotionally responsive AI intervention is more associated with decreasing speaking anxiety and increasing engagement compared to an emotion-agnostic intervention, but that these benefits are more pronounced when emotional support is also accompanied by real pedagogical and cognitive adaptation, and when novelty-driven interest is accompanied by meaningful gains in learner competence. Meanwhile, there are clear boundary conditions to sense in the literature: emotion-assessment technologies need to be further cross-culturally validated; teachers'–learners' interactions remain an important influence on enjoyment and anxiety, and the collection of biometric data has privacy and equity concerns that must be actively managed. Future work should focus on longitudinal, classroom-integrated assessment of the type of integrated emotion-aware system suggested here, especially regarding the evaluation of the fairness of classifiers across different student cohorts, and the distinction of enduring effects of engagement from effects due to novelty. Under these conditions, emotion-aware AI has potential to make ELL more affective responsive.

References

- Altuwairqi, K., Jarraya, S. K., Allinjawi, A., & Hammami, M. (2021). Student behavior analysis to measure engagement levels in online learning environments. *Signal, image and video processing*, 15(7), 1387-1395. <https://doi.org/10.1007/s11760-021-01869-7>
- Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of learning analytics*, 3(2), 220-238. <https://doi.org/10.18608/jla.2016.32.11>
- Chandel, P., & Lim, F. V. (2024). Generative AI and literacy development in the language classroom: A systematic review of literature. *Ubiquitous Learning*, 18(2), 31. <https://doi.org/10.18848/1835-9795/CGP/v18i02/31-49>
- Dewaele, J. M., & Alfawzan, M. (2018). Does the effect of enjoyment outweigh that of anxiety in foreign language performance? *Studies in second language learning and teaching*, 8(1), 21-45. <https://doi.org/10.14746/ssllt.2018.8.1.2>
- Dewaele, J. M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in second language learning and teaching*, 4(2), 237-274. <https://doi.org/10.14746/ssllt.2014.4.2.5>
- Dewaele, J. M., Albakistani, A., & Ahmed, I. K. (2024). Levels of foreign language enjoyment, anxiety and boredom in emergency remote teaching and in in-person classes. *The Language Learning Journal*, 52(1), 117-130. <https://doi.org/10.1080/09571736.2022.2110607>
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in human behavior*, 75, 461-468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in human Behavior*, 93, 279-289. <https://doi.org/10.1016/j.chb.2018.12.023>
- Liu, Y., Zhang, H., Jiang, M., Chen, J., & Wang, M. (2024). A systematic review of research on emotional artificial intelligence in English language education. *System*, 126, 103478. <https://doi.org/10.1016/j.system.2024.103478>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11(1), 50. <https://doi.org/10.1186/s40561-024-00342-5>
- Monkaresi, H., Bosch, N., Calvo, R. A., & D'Mello, S. K. (2016). Automated detection of engagement using video-based estimation of facial expressions and heart rate. *IEEE Transactions on Affective Computing*, 8(1), 15-28. <https://doi.org/10.1109/TAFFC.2016.2515084>
- Peña-Acuña, B., & Corga Fernandes Durão, R. (2024, November). Learning English as a second language with artificial intelligence for prospective teachers: a systematic review. In *Frontiers in education* (Vol. 9, p. 1490067). Frontiers Media SA. <https://doi.org/10.3389/feduc.2024.1490067>
- Picard, R. W. (2003). Affective computing: challenges. *International journal of human-computer studies*, 59(1-2), 55-64. [https://doi.org/10.1016/S1071-5819\(03\)00052-1](https://doi.org/10.1016/S1071-5819(03)00052-1)
- Rojas Vistorte, A. O., Deroncele-Acosta, A., Martín Ayala, J. L., Barrasa, A., López-Granero, C., & Martí-González, M. (2024). Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review. *Frontiers in psychology*, 15. <http://doi.org/10.3389/fpsyg.2024.1387089>

- Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System*, 127, 103533. <https://doi.org/10.1016/j.system.2024.103533>
- Whitehill, J., Serpell, Z., Lin, Y. C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing*, 5(1), 86-98. <https://doi.org/10.1109/TAFFC.2014.2316163>