

Research Article:

AI-Supported Language Learning for Non-Native English-Speaking Students: An Action Research

Vinothini Vasodavan* and Jasmine Jain

School of Education, Faculty of Social Sciences and Leisure Management, Taylor's University, 47500 Subang Jaya, Selangor, Malaysia

*Corresponding author: vino.vasodavan@taylors.edu.my

ABSTRACT

As linguistic diversity in higher education expands, non-native English-speaking students face persistent language barriers that hinder engagement, learning outcomes, and equity. This study investigates the impact of generative artificial intelligence (Gen-AI) in addressing these challenges through Walter, an AI-powered learning buddy designed to provide personalised linguistic and academic support. Framed by Sociocultural Theory, Walter functioned as a More Knowledgeable Other (MKO), scaffolding learners within their Zone of Proximal Development (ZPD) to enhance proficiency and foster academic participation. A mixed methods design integrated quantitative usage analytics and oral presentation performance ($n = 30$) with qualitative thematic analysis of student and instructor perspectives. Over a six-week intervention, analytics showed distinct patterns of in-campus and off-campus engagement, with usage peaks corresponding to high-stakes assessment preparation. Paired-sample t -tests revealed a statistically significant improvement in integrated oral presentation scores ($p < .001$), with a large effect size ($d = 0.75$) indicating substantial pedagogical impact. Qualitative findings identified three key themes: Walter as a psychologically safe space for rehearsal, its adaptive integration into diverse learning routines, and recognised limitations regarding conversational depth. While findings highlight Walter's capacity to build skill and learner confidence, challenges regarding cultural sensitivity and contextual relevance remain. This study contributes empirical evidence for the role of Gen-AI in fostering inclusive higher education, recommending that institutions move beyond general-purpose AI toward discipline-specific, strategically integrated frameworks to ensure equitable and effective learning environments for all students.

Keywords: AI learning buddy, language barrier, higher education, Gen-AI, non-native English-speaking

Published: 9 June 2026

To cite this article: Vasodavan, V., & Jain, J. (2026). AI-supported language learning for non-native English-speaking students: An action research. *Asia Pacific Journal of Educators and Education*, 41(1), 319–336. <https://doi.org/10.21315/apjee2026.41.1.14>

INTRODUCTION

The globalisation of higher education has led to an increased presence of international students, particularly from non-English-speaking countries such as China. These students often encounter significant challenges in academic English, especially when engaging in writing, comprehension of lectures, and classroom participation (Wei et al., 2024; Heng, 2018). Despite two decades of pedagogical effort, recent literature confirms that non-native English-speaking (NNES) students still face pervasive linguistic, cognitive, and affective barriers in higher education. Specifically, contemporary research highlights that limited vocabulary and L2 knowledge continue to hinder academic comprehension (Shah et al., 2025), while language anxiety and fear of judgement lead to social isolation and disengagement (Simanjuntak & Simanjuntak, 2025; Okoli & Nweke, 2024). These findings collectively emphasise that effective language support must address both the academic skill deficit and the underlying psychological safety concerns to foster equitable learning outcomes.

Recent advancements in educational technology have positioned artificial intelligence (AI) as a potential solution to support personalised learning at scale. Conversational AI, like chatbots and AI-powered learning assistants, has gained attention for its ability to provide immediate feedback, adaptive language practice, and interactive communication experiences (Aliakbari et al., 2025; Guan et al., 2025). AI tools can scaffold vocabulary acquisition, grammar usage, and communicative competence by simulating real-life dialogues and promoting sustained learner engagement (Holmes et al., 2021). This is particularly relevant in higher education contexts where students must navigate self-directed and content-heavy curricula in a non-native language.

However, a critical challenge has emerged within the global higher education context: while AI tools are widely recognised for providing a psychologically safe space that reduces speaking anxiety, their ability to produce measurable proficiency gains remains inconsistent (Yunita et al., 2025). Global evidence suggests that while AI shows moderate success in vocabulary acquisition (Cohen's $d = 0.61$) and writing revision, it often fails to significantly improve real-time oral fluency compared to traditional instruction (Ramli et al., 2025; Yunita et al., 2025). Furthermore, technical limitations, such as a documented 23% drop in speech recognition accuracy for non-native accents, often transform these potential facilitators into barriers for international learners (Ramli et al., 2025). There is, therefore, a pressing need to evaluate whether structured, pedagogical AI agents can bridge this affective-linguistic gap by moving beyond mere affective comfort to facilitating demonstrable skill acquisition (Yunita et al., 2025).

Despite the proliferation of general-purpose AI tools like ChatGPT, there is a significant lack of empirical research specifically examining dedicated remedial AI agents, such as Walter, within structured university frameworks (Oktarina et al., 2024). Current global literature is characterised by a temporal disconnect where affective benefits such as increased confidence are reported immediately, yet the acquisition of cognitive linguistic

skills requires more sustained, longitudinal evidence (Yunita et al., 2025). This study addresses these gaps by evaluating the use of Walter, an AI Learning Buddy, within the WALT-Buddy remedial programme. By investigating the trajectory from reduced learner anxiety to measurable proficiency improvements, this research contributes to emerging scholarship on the pedagogical potential of AI in authentic, digitally mediated higher education environments.

This study investigates the use and impact of Walter, the AI Learning Buddy, by addressing the following research questions:

RQ1: In what ways do non-native English-speaking students use Walter, the AI Learning Buddy, during the intervention period in the WALT-Buddy remedial English language programme?

RQ2: To what extent does the use of Walter, the AI Learning Buddy, contribute to improvements in English language proficiency among non-native English-speaking students, as measured by pre- and post-intervention assessments from the WALT-Buddy remedial English language programme?

RQ3: How do non-native English-speaking students perceive their experience of using Walter, the AI Learning Buddy, in supporting their English language development throughout the WALT-Buddy remedial English language programme?

LITERATURE REVIEW

AI in Language Learning: Evidence, Affordances, and Gaps

The international landscape of higher education is currently defined by a fundamental paradigm shift from rigid, rule-based Computer-Assisted Language Learning (CALL) to fluid, generative AI and Large Language Models (LLMs) (Albedah, 2025; Dou et al., 2025). This evolution represents more than a technical upgrade as it is a transition from static drills to contemporary AI-driven environments that function as collaborative, individualised partners (Dharani & Shobana, 2025; Zhu & Wang, 2025). This global transformation is increasingly led by research output from Asian nations, which now dominate the bibliometric landscape (Alaqlobi et al., 2024; Han, 2026).

Empirical evidence across the literature substantiates that this shift yields significant, measurable proficiency gains. Meta-analytic data reveals substantial effect sizes for AI-enhanced instruction, particularly in speaking proficiency ($SMD = 1.033$) and linguistic development ($d = 1.18$) (Jantakoon et al., 2025; Lee & Lee, 2024). Furthermore, the literature highlights a critical distinction between generic and tailored tools; culturally specific AI systems achieve significantly larger vocabulary gains ($d = 0.85$) than their generic counterparts (Ramli et al., 2025). However, a tension exists in the data, while writing and speaking gains are robust, listening outcomes often fail to reach statistical

significance, suggesting that AI's affordances are currently skill-dependent rather than universal (Jantakoon et al., 2025).

Beyond cognitive gains, the most consistent affordance identified in recent research is AI's capacity to mitigate Foreign Language Anxiety (FLA). AI platforms provide a psychologically safe, non-judgemental environment that is particularly vital in face-saving cultures where the silent classroom phenomenon historically inhibits participation (Yunita et al., 2025; Ma & Chen, 2025). Learners demonstrate a higher willingness to communicate (WTC) with objective machines than with human peers (Fathi et al., 2024; Zhi & Wang, 2024). Yet, a critical gap remains as the current literature struggles to bridge the divide between these immediate affective benefits (confidence) and long-term cognitive mastery, which often requires interventions longer than the typical eight-week study period (Han, 2026; Xu & Wang, 2024). This disconnect justifies the need for the WALT-Buddy program, which seeks to integrate these affective safe spaces into a structured pedagogical framework.

Language Barriers and Institutional Gaps in Higher Education

The challenges facing NNES students extend beyond simple linguistic deficits to encompass severe time-costs and cultural barriers. NNES scientists spend significantly more time reading and writing than native speakers (Amano et al., 2023), a burden compounded for Chinese international students by cultural norms like face-saving and respect for authority (Nguyen, 2020; Yi & Meng, 2022). While traditional institutional support such as writing centres attempts to mitigate these issues, the literature critiques these models for being one-size-fits-all and failing to scale to meet diverse student needs (Hyland, 2019; Naidoo, 2017).

Furthermore, a readiness gap persists where student enthusiasm for AI outpaces institutional infrastructure and faculty training (Tong & Tran, 2024). This discrepancy is often exacerbated by cultural stigma surrounding help-seeking behaviour (Jack et al., 2019). When educators assume cultural similarity, communication breakdowns occur, highlighting a profound need for mediators that can bridge these norms (Dumessa & Godesso, 2013). This institutional failure creates a specific niche for the WALT-Buddy, a tool that provides the human touch of empathy and cultural nuance that automated tools often lack (Anwar et al., 2025), serving as a blended bridge between the student and the institution.

Theoretical Framework: Sociocultural Learning and Technological Mediation

This study synthesises Sociocultural Learning Theory (SCL) and Technological Mediation Theory to conceptualise the AI Learning Buddy as a non-human communicative actor. Following Vygotsky (1978), learning is viewed as a socially situated process occurring within the Zone of Proximal Development (ZPD). Within this framework, Walter functions as a mediational artefact (Ridha et al., 2026). By offering dialogic prompts,

Walter facilitates the transition from other regulation to independent self-correction (Qiao & Zhao, 2023). Critically, this extends Verbeek's (2015) Technological Mediation Theory, which argues that such technology is not a neutral tool but a quasi-other that transforms human perception.

This relational ontology is particularly relevant for Chinese international students, as it allows for a lowering of the affective filter necessary for language internalisation (Aagaard & Rosenberger, 2024; Yunita et al., 2025). However, the literature warns of an AI Paradox where helpful feedback can lead to automation complacency and the erosion of authentic identity (Seddiki & Korichi, 2026). Consequently, this study posits that the WALT-Buddy is most effective not as a replacement for the learner, but as a mediator within a distributed communication model where student agency remains central (Farber, 2025; Long et al., 2025).

METHODS

This study employed a qualitative-dominant mixed-methods action research design to investigate the use of an AI-powered learning companion (Walter) to support foreign language learners in improving their English language proficiency under a programme known as Walter's Language Buddy Program (WALT-Buddy) Action research was selected due to its iterative, participatory, and practice-oriented nature, enabling the researcher to implement, evaluate, and refine the intervention in a real educational setting (Kemmis et al., 2014). The study focused on understanding not only the effects of the intervention but also the students' engagement and experiences throughout the learning process.

Intervention: The WALT-Buddy

The intervention was conducted at a private university in Malaysia, situated within a six-week WALT-Buddy remedial English language programme designed for international undergraduate students. The study utilised a purposive sampling technique, selecting a unique and academically critical cohort of 30 students ($n = 30$) who met the criteria for intensive language support. These participants were predominantly Mandarin-dominant, non-native English speakers enrolled in a specialised undergraduate Education programme. This specific sample size was a necessary constraint dictated by the intersection of the students' linguistic profiles, documented low proficiency, and their professional training context. Accessing such a specialised group is inherently challenging; therefore, the 30 participants represented the maximum practical recruitment for this phase of the study. All participants consented to the research, offering a meaningful context for investigating the challenges of adapting to English-medium instruction within internationalised Malaysian higher education and the subsequent impact of AI-mediated intervention.

As an intervention programme, WALT-Buddy is designed to produce proximal, rapid gains in fundamental skills necessary for subsequent academic success. The 6-week period was selected as a sufficient duration for measuring the initial learning acceleration required for remedial cohorts. The WALT-Buddy intervention involved the integration of Walter,

an AI-powered chatbot powered by the Noodle Factory (NF) platform, into the students' self-paced and instructor-led activities.

Walter was specifically engineered to provide real-time, text-based interaction, offering features such as personalised vocabulary feedback, grammar support, paraphrasing assistance, and conversational simulation. Students were introduced to the tool in Week 1 and engaged in regular independent practice and task-based assignments. Each week, language activities were designed to integrate Walter into writing, reading comprehension, and speaking support based on the themes presented in Table 1.

Table 1. Themes of student activities each week during the WALT-Buddy intervention

Week	Students' activities on campus	Students' activities off campus
1	Onboarding session on the WALT-Buddy and Diagnostic Presentation (Pre-test). Students deliver a short presentation (2 minutes) about themselves and their culture to the class.	Self-directed on how the platform's tools, like quizzes and grammar checkers, will be used to complete both on-campus and off-campus activities throughout the programme.
2	Topic 1: Hello, I'm a Future Teacher. This session focuses on basic expression. This session focus on pronunciation and reading (pronunciation) exercises.	Use voice features in the NF platform to practice expressions used by teachers. Audio recording 50 expressions on presenting oneself in 60 seconds.
3	Topic 3: Mission Classroom Commander: Peer Role-Playing. In pairs, students take turns acting as the teacher and student, giving and following a set of instructions.	Create and record a 2-minute dialogue after role play with WALT-Buddy simulating a teacher-student interaction.
3	Topic 4: Detective Reading Group & Think-Pair-Share: In small groups, students discuss the article. They share key vocabulary and main ideas.	Use WALT-Buddy's automated quizzes to practice and assess their reading comprehension and vocabulary.
4	Topic 5: The Teacher's Toolkit Writing. In a small group, students work together to draft a simple presentation about a new teaching method.	Use WALT-Buddy's essay rubric checker to refine and edit the presentation to fit the given topics.
5	Topic 6: The Pitch Perfect Challenge. Students work in small groups to collaboratively design a 2-minute to elevator pitch for a new teaching method.	Use WALT-Buddy's contextual chat to brainstorm content and talking points for the pitch and explore different design ideas for their presentation slides.
6	Presentation (post-test). Students deliver a short presentation (2 minutes) about themselves and their culture to the class.	

The methodological justification for attributing language gains to this intervention lies in Walter's capacity for instant scaffolding. Unlike traditional classroom settings, where teacher feedback is often delayed, Walter provided immediate corrective feedback loops during independent study. This allowed for repetitive, low-stakes practice that directly addressed individual linguistic gaps as they occurred. While Walter did not replace the instructor, it served as a continuous supplementary tutor available beyond classroom hours, significantly increasing students' time-on-task and providing a personalised learning environment that is difficult to achieve in standard group-based remedial instruction. Usage data were automatically logged to monitor intervention fidelity and student engagement throughout the 6-week period.

DATA ANALYSIS

Quantitative Data Analysis

Quantitative data were utilised to address Research Objectives 1 (Usage Patterns) and 2 (Language Proficiency Outcomes). To evaluate usage patterns (RO1), system-generated logs were extracted from the Noodle Factory platform. These logs captured granular data, including the frequency, duration, and specific types of student interactions with the AI agent, Walter. Descriptive statistical analysis was then performed to examine engagement trends and ensure intervention fidelity across the six-week period. To measure language proficiency outcomes (RO2), a pre-test and post-test design was implemented via an integrated oral presentation task. This assessment was designed through the lens of Culturally Responsive Pedagogy (CRP), allowing students to choose from four culturally relevant themes: *International Learning Experiences*, *My Hometown Heritage*, *Future Career Plans*, or *Cultural Diversity*. This approach was intended to foster deeper engagement and comfort, particularly for student-teachers preparing for diverse classroom environments.

The assessment utilised a holistic rubric where the four core language skills were equally weighted at 25% each. Speaking was evaluated based on delivery flow, coherence, and confidence during the presentation, while writing was measured through the clarity, grammar, and vocabulary precision of the content authored on the presentation slides. Reading proficiency was assessed by the students' ability to synthesise information from two academic texts related to their chosen topic to inform their slide content, and listening was evaluated through a mandatory three-minute spontaneous Q&A session with the assessor. Statistical analysis for RO2 was conducted using paired-sample *t*-tests ($n = 30$) to determine the significance of performance gains between the pre-test and post-test phases. Finally, effect sizes were calculated using Cohen's *d* to determine the practical magnitude of the intervention's impact on the students' overall language proficiency.

Qualitative Data Analysis

Qualitative data were collected to address Research Objective 3 (Learner Experiences), focusing on participants' perceptions of the utility and limitations of the Walter AI agent. Semi-structured interviews were conducted at the conclusion of the WALT-Buddy intervention with a purposive subsample of six students. These interviews explored the nuances of participant experiences, including the specific contexts and frequency of Walter's use both on and off campus, as well as the integration of supplementary instructional tools, such as the use of ChatGPT during on-campus brainstorming sessions. To ensure a comprehensive understanding of the programme's impact, the analysis also incorporated reflective feedback from the four instructors involved in the delivery of the remedial programme.

The qualitative data were transcribed verbatim and analysed using Braun and Clarke's (2006) six-phase thematic analysis framework. This rigorous process involved initial data familiarisation, the generation of codes, and the eventual development and refinement of themes that captured the essence of the learners' and instructors' experiences. By triangulating student interview responses with instructor reflections, the analysis aimed to provide a multi-perspective view of the intervention's effectiveness. This thematic approach allowed for the identification of meaningful patterns regarding how AI-mediated scaffolding supports remedial English language development in an internationalised higher education context.

FINDINGS

The findings will be reported according to its corresponding research questions:

RQ1: In what ways do non-native English-speaking students use Walter, the AI Learning Buddy, during the intervention period in the WALT-Buddy remedial English language programme?

Usage analytics from the NF dashboard revealed a total of 27 chat sessions with a mean of 0.47 daily sessions. While the average engagement was sporadic, the longitudinal distribution (Figure 1) showed two significant peaks in total minutes corresponding directly to the mid-way formative assessment and the final summative oral presentation. This pattern suggests that students utilised Walter primarily as a just-in-time scaffold for high-stakes tasks rather than a consistent longitudinal companion. While this could be interpreted as low engagement, an alternative interpretation suggests a high level of strategic agency. From a Sociocultural (SCL) perspective, the learners activated the AI mediator only when their ZPD was challenged by external academic pressure.

Furthermore, the maximum conversational turn count of 14 indicates that while many sessions were brief, a subset of students engaged in sustained linguistic negotiation. This creates a theoretical implication for remedial NNES students, the quasi-other (Verbeek,

2015) is not a constant presence but a situational partner that is called upon to mitigate risk during periods of high cognitive load. This complicates the ideal of AI as a 24/7 tutor, suggesting instead that its primary value lies in its on-demand availability during academic crises.

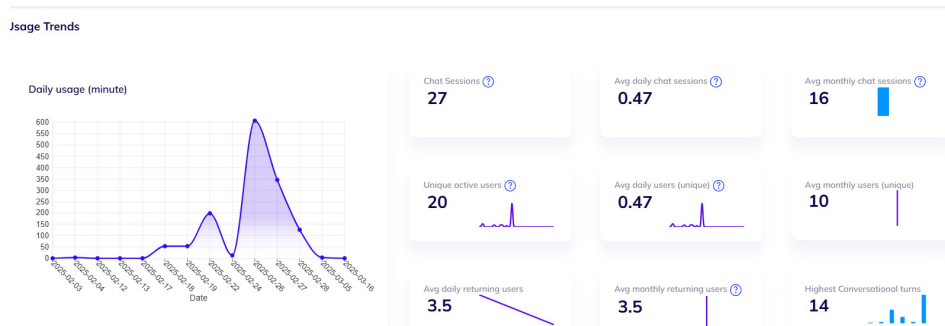


Figure 1. Six-week usage analytics of Walter from the Noodle Factory (NF) dashboard

RQ2: To what extent does the use of Walter, the AI Learning Buddy, contribute to improvements in English language proficiency among non-native English-speaking students, as measured by pre- and post-intervention assessments from WALT-Buddy remedial English language programme?

To address this question, a paired-sample t -test was conducted to compare students' pre-test and post-test scores on an oral presentation assessment that integrated all four components of English language proficiency: speaking, listening, reading, and writing (see Table 2). A statistically significant improvement was found in oral presentation scores from pre-test ($M = 61.15$, $SD = 10.80$) to post-test ($M = 69.40$, $SD = 11.20$), $t(29) = 4.11$, $p < .001$, with a very large effect size ($d = 0.75$). The resulting effect size ($d = 0.75$) is categorised as moderate-to-large, indicating consistent and substantial gains across the participant cohort. These results suggest that Walter was associated with marked improvements in integrated language performance. However, a deeper interrogation of this moderate-to-large effect is necessary to distinguish between linguistic acquisition and performance scaffolding. One alternative interpretation is that the score increase represents a reduction in the affective filter (Krashen, 1982) rather than a fundamental shift in cognitive linguistic competence. If Walter functioned primarily as a psychological safety net, the $d = 0.75$ gain may reflect students' increased willingness to deploy existing knowledge more confidently, rather than the acquisition of new grammatical structures.

Furthermore, the 95% confidence interval (4.15 to 12.35) suggests that while gains were universal, they were also varied. This leads to the theoretical implication that AI-mediated scaffolding (Vygotsky, 1978) is highly effective at bridging the gap for integrated tasks like public speaking, but its role as a permanent driver of proficiency requires longitudinal verification. By moving beyond rote learning toward holistic communicative confidence,

the WALT-Buddy programme appears to facilitate a curriculum of wholeness, yet the extent to which these gains are independent of the AI mediator remains a critical question for future research.

Table 2. Paired-sample *t*-test for oral presentation scores incorporating speaking, listening, reading, and writing.

Measure	Pre-test <i>M</i>	SD	Post-test <i>M</i>	SD	<i>t</i>	df	<i>p</i>	Cohen's <i>d</i>	95% CI [LL, UL]
Oral presentation score	61.15	10.80	69.40	11.20	4.11	29	.001	0.75	4.15, 12.35

Note: Values represent mean scores (*M*) and standard deviations (*SD*) for oral presentation assessments before and after the six-week intervention. Cohen's *d* was calculated for paired samples. CI = confidence interval; LL = lower limit; UL = upper limit.

RQ3: How do non-native English-speaking students perceive their experience of using Walter, the AI Learning Buddy, in supporting their English language development throughout the WALT-Buddy remedial English language programme?

Thematic analysis of the semi-structured interviews was conducted using Braun and Clarke's (2006) six-phase framework, a process that moved beyond descriptive coding to an interpretive mapping of student experiences against the study's theoretical foundations. This analysis was strengthened through data source triangulation, where student perspectives were cross-checked with qualitative contextual usage accounts of in-campus and off-campus engagement, insights on supplementary tool integration (e.g., ChatGPT), and reflective notes from the four programme instructors. By viewing the data through the lenses of sociocultural learning and technological mediation, the analysis identified three interconnected themes. These themes capture how Walter functioned as a scaffolded mediator of learner experience and how the development of trust in AI tools shaped the day-to-day realities of the WALT-Buddy remedial English language programme.

Theme 1: Perceived Psychological Safety and Autonomy in Language Practice

The analysis revealed that students perceived Walter as a low-stakes environment, which fundamentally identifies the AI agent as a primary tool for technological mediation. Rather than a simple utility, Walter functioned through an alterity relation (Verbeek, 2015), where the machine is perceived as a quasi-other. This specific relationship allowed learners to bypass the face-saving sociocultural pressures prevalent in traditional classrooms. The interpretive value here lies not just in the lack of shyness, but in the deliberate outsourcing of the judgemental gaze to a non-human entity. This mediation transformed the learner experience by providing a safe digital space that lowered the affective filter (Krashen, 1982), as evidenced by the participants:

It feels like having a patient tutor who never gets tired of answering my questions, even if I ask them again and again. (Aisha, L42–L44)

When I talk to Walter, I can try my English without feeling shy because no one is watching me. (Mei Ling, L58–L60)

Sometimes I practise my sentences before I say them in class, so I feel more confident. (Rahim, L73–L75)

Aisha's reflection on patience suggests that the mediation is as much about time as it is about tone; the AI removes the social guilt associated with slow progress. This perception reflects principles of psychological safety, where learners feel secure enough to test new linguistic structures without fear of social embarrassment. The development of trust in the AI tool was therefore not purely technical, but relational; students trusted the agent's patience to facilitate the transition from internal rehearsal to public performance.

Theme 2: Adaptive Use of Walter Across Learning Contexts

The analysis revealed that students' perceptions of utility were contingent upon the context of engagement, illustrating a sophisticated movement through the ZPD. This suggests that the Learning Buddy is not a static scaffold but a fluid mediator that expands or contracts based on the physical environment. In-campus use functioned as a co-scaffold alongside human instructors, whereas off-campus use saw Walter transition into the primary More Knowledgeable Other (MKO). This shift indicates a high level of learner autonomy, where students consciously decided when to rely on a human and when to rely on the quasi-other for validation.

I usually use Walter in class when the teacher asks us to brainstorm ideas together. (Farah, L91–L93)

At home, I use Walter the most. It helps me practise my presentation many times until I feel ready. (Daniel, L107–L109)

Sometimes I open NF in the class to check pronunciation before I speak in front of the group. (Sofia, L120–L122)

Daniel's use of Walter to practice until I feel ready highlights the AI's role in pre-scaffolding confidence before the student enters a social-evaluative space. The strategic pairing of Walter with other tools like ChatGPT suggests that students were not merely passive recipients of AI; they were active architects of a personalised, technologically mediated learning ecosystem.

Theme 3: Limitations in Practice and Suggestions for Development

A critical interpretive finding is that trust in the AI tool was not a permanent state, but an evaluative process that shifted as the student's own competence grew. As students moved through the ZPD, their requirements for mediation changed from basic accuracy to social authenticity. This evaluative agency suggests that the more proficient a student becomes, the more they interrogate the machine-like nature of the AI.

Sometimes the answers feel too general, so I still need to ask my teacher for more details. (Ali, L138–L140)

I want Walter to sound more natural, like having a real conversation, not just giving short answers. (Mei Ling, L151–L153)

I want more listening practice with audio so I can hear the correct pronunciation. (Rahim, L165–L167)

Mei Ling's desire for natural conversation indicates a shift in the perceived ZPD as she no longer needs a simple correction tool, but a sophisticated social interlocutor. This highlights a theoretical tension while AI provides a psychological safe haven. These generic linguistic boundaries can eventually become a barrier to the communicative competence that required for real-world interaction. This feedback serves as evidence of the students' active role in their own development, demonstrating that their agency eventually outgrows the initial scaffold provided by the WALT-Buddy.

DISCUSSION

The results of this study provide empirical evidence that Walter, the AI Learning Buddy, appears to function as a technological mediator within the ZPD. The statistically significant improvement in oral presentation scores $t(29) = 4.11, p < .001$ with a moderate-to-large effect size ($d = 0.75$) suggests that the intervention may have helped bridged the gap between learners' current abilities and their potential linguistic performance. These findings align with Dharani and Shobana (2025), who posit that AI-integrated environments accelerate language acquisition through structured, iterative feedback.

Reducing the Affective Filter through AI Mediation

A primary finding of this study is the reduction of language anxiety, which served as a catalyst for the observed quantitative gains. Theme 1 (Psychological Safety) reveals that students perceived Walter as a low-stakes interlocutor, a sentiment that correlates with the significant score increases. This supports the recent findings of Li et al. (2025) and Yunita et al. (2025), who argue that Gen-AI tools effectively lower the affective filter by providing a non-judgemental practice space. However, this finding warrants a deeper interrogation of the nature of these gains. By acting as a safe mediator, Walter allowed students to engage

in linguistic risk-taking and rehearsing complex structures before facing a human audience. An alternative interpretation suggests that the $d = 0.75$ effect size is not merely a result of increased practice time, but a direct outcome of the psychological safety inherent in AI-mediated pedagogies, which unlocks existing knowledge that was previously suppressed by the silent classroom phenomenon (Nguyen, 2020).

Situating Usage Patterns in Sociocultural Scaffolding

The usage analytics revealed concentrated peaks of engagement such as 3.5 returning users daily particularly surrounding high-stakes assessments. This usage pattern confirms the role of Walter as an adaptive scaffold that students strategically utilised for preparation and pronunciation checking. These results resonate with Nadeem (2025), who noted that self-directed AI use often follows a just-in-time logic, where learners leverage the tool's 24/7 accessibility to supplement classroom instruction. However, this just-in-time use creates a conflict if students only use the AI as a safety net during exams, it functions more like a rescue tool than a long-term learning partner. This suggests that while Walter is great for reducing immediate stress, he may not yet be integrated enough into the students' daily study habits to create lasting language change. Furthermore, the qualitative findings regarding generic responses highlight a ceiling effect in AI scaffolding. While AI provides breadth of access, it may still struggle with the communicative competence and cultural nuance required for advanced social situatedness (Hong, 2023). This indicates that as learners' proficiency grows and they move toward the upper limits of the ZPD, their requirements for technological mediation become more complex, shifting from a need for basic accuracy to a demand for social authenticity.

Learner Agency and the Paradox of Trust

Despite the robust statistical gains, learners demonstrated significant agency in evaluating Walter's limitations. Many students transitioned from initial curiosity to critical assessment, seeking more natural conversational loops; this finding supports Jantakoon et al. (2025) regarding the need for human-like fluidity in educational AI. The observed anthropomorphism where students described Walter as a patient tutor suggests that the AI's social presence was a key driver of engagement. However, this creates a dual-edged risk if students view the AI as an infallible, human-like authority, they may lose the critical distance needed to evaluate the accuracy of the AI's feedback. In such cases, students might prioritise 'algorithmic approval' over their authentic linguistic identity (Seddiki & Korichi, 2026). These contradictions underscore the importance of integrating AI not as a standalone solution, but as a critical component of a curriculum of wholeness that prioritises both technical proficiency and the social dimensions of learning.

IMPLICATIONS AND CONCLUSION

The findings of this study offer several pedagogical implications for the integration of AI in language education. Primarily, the results demonstrate that Walter, the AI Learning Buddy,

serves as an effective tool for technological mediation, providing a low-stakes environment that significantly lowers the affective filter for remedial learners. The deeper implication for practitioners is that AI tools are most effective when framed not as a replacement for human instruction, but as a transitional scaffold within a sociocultural framework. By offering a private, non-judgemental space for rehearsal, Walter addresses the psychological barriers such as performance anxiety that often hinder non-native speakers. This suggests that the quasi-other status of the AI as is its greatest pedagogical asset, allowing students to bridge the gap between silent observation and active communicative confidence (Verbeek, 2015).

Regarding institutional policy, the empirical evidence from the WALT-Buddy programme provides a preliminary basis for exploring broader AI adoption in higher education. While the study suggests that AI-driven tools can help mitigate resource constraints, the researchers advocate for a proportionate and incremental policy shift rather than a high-scale mandate. Institutions should consider a pilot-based integration of AI scaffolds within existing remedial frameworks, prioritising the development of discipline-specific AI mediation. This ensures that technological support is tailored to the specific academic jargon and scenarios relevant to the students' fields of study. Furthermore, funding should be directed toward establishing a digital infrastructure for equity, ensuring that the benefits of AI-mediated practice are not limited by a student's initial proficiency or socioeconomic status.

In conclusion, this action research project provides valuable, context-specific insights into the role of generative AI in supporting international student success. However, a critical interrogation of the moderate-to-large effect size ($d = 0.75$) reminds us that these results should be viewed as exploratory evidence rather than a definitive long-term proof of efficacy. This necessitates future longitudinal research to assess linguistic retention and the impact of AI scaffolding over 6-month or 1-year intervals. Nevertheless, through the triangulation of quantitative analytics and qualitative themes, this study contributes a rigorous and balanced perspective to the growing body of research on AI-supported language learning. It highlights the potential for AI to foster a more inclusive curriculum of wholeness that recognises the learner as an active agent in a technologically mediated world.

LIMITATIONS OF THE STUDY

While this study provides valuable insights into the technological mediation of AI in remedial language contexts, several limitations must be acknowledged. First, the small sample size and the focus on a specific cohort of international students at a single institution limit the generalisability of the findings. As an action research project, the primary goal was localised instructional improvement; therefore, the results should be interpreted as exploratory evidence rather than a universal model for AI integration. Second, the six-week intervention period, while sufficient to show immediate gains in oral presentation scores, does not account for long-term linguistic retention or the potential fading of the AI scaffold over time.

Furthermore, the qualitative data relied on semi-structured interviews and instructor reflections, which may be susceptible to social desirability bias, where participants provide responses, they believe align with the researchers' expectations. Finally, while the triangulation of usage analytics and qualitative themes strengthens the study's internal validity, the quantitative analysis did not include a control group. Consequently, while the statistically significant improvement is encouraging, it cannot be solely attributed to the AI tool without accounting for other confounding variables in the students' academic environment. Future longitudinal research involving larger, more diverse populations and a quasi-experimental design is recommended to further validate the long-term efficacy of AI-scaffolded pedagogies.

ETHICAL CONSIDERATIONS

Ethical approval was obtained from the university's research ethics committee. All participants were informed of the voluntary nature of the study, data confidentiality, and their right to withdraw at any point. Pseudonyms were used in all reporting to ensure anonymity. Informed consent was secured in both written and verbal forms at the outset of the study.

DATA AVAILABILITY

The datasets generated during the current study are not publicly available as the authors do not wish to share the data as this may breach participant confidentiality but are available from the corresponding author on reasonable request.

REFERENCES

- Aagaard, J., & Rosenberger, R. (2024). Experience without essentialism: On 'posting' phenomenology. *Qualitative Research in Psychology*, 21(1), 1–19. <https://doi.org/10.1080/14780887.2023.2294116>
- Alaqlobi, O., Alduais, A., Qasem, F., & Alasmari, M. (2024). Artificial intelligence in applied linguistics: A content analysis and future prospects. *Cogent Arts & Humanities*, 11(1), 2382422. <https://doi.org/10.1080/23311983.2024.2382422>
- Albedah, F. (2025). Artificial intelligence in language education: A systematic review of multilingual applications, large language models, and emerging challenges. *Language Teaching Research Quarterly*, 49, 247–268. <https://doi.org/10.32038/ltrq.2025.49.04>.
- Aliakbari, M., Barzan, P., & Sayyadi, M. (2025). Exploring the impact of AI chatbots on EFL learners' conversational proficiency. *Journal of Interdisciplinary Research in English Language Communication*, 1(2), 66–80. <https://doi.org/10.30470/IRELC.2025.2058800.1022>

- Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M., González-Trujillo, J. D., Montaña-Centellas, F., Paudel, K., White, R. L., & Verissimo, D. (2023). The manifold costs of being a non-native English speaker in science. *PLOS Biology*, 21(7), e3002184. <https://doi.org/10.1371/journal.pbio.3002184>
- Anwar, Q., Shabbir, Z., Raza, A., & Baloch, C. (2025). Using artificial intelligence to enhance language acquisition within educational settings. *Dialogue Social Science Review*, 6(1), 112–125.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Dharani, K., & Shobana, R. (2025). The impact of AI-driven language models on second language acquisition. *International Journal of English Language and Literature Research*, 3(1), 10–18. <https://doi.org/10.56472/25842773/IJELLR-V3I1P103>
- Dou, A., Xu, W., Li, X., Zhang, S., & Zhang, J. (2025). Artificial intelligence in language learning: A systematic review and bibliometric analysis. *International Journal of Distance Education Technologies*, 23(1), 1–24. <https://doi.org/10.4018/IJDET.385045>.
- Dumessa, M., & Godesso, M. (2013). Interethnic relations among students of Jimma University, Oromiya, Ethiopia. *International Journal of Sociology and Anthropology*, 5(4), 119–128. <https://doi.org/10.5897/IJSA2013.0446>
- Farber, B., & Smith, J. (2025). Comparing human and AI expertise in the academic peer review process: Towards a hybrid approach. *Higher Education Research & Development*, 44(2), 1–15. <https://doi.org/10.1080/07294360.2024.2445575>
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254>
- Guan, R., Raković, M., Chen, G., & Gašević, D. (2025). How educational chatbots support self-regulated learning? A systematic review of the literature. *Education and Information Technologies*, 30(4), 4493–4518. <https://doi.org/10.1007/s10639-024-12881-y>
- Han, Y. (2026). Envisioning the futures of language education in the era of artificial intelligence. *Journal of Futures Studies*, 30(2), 15–34.
- Heng, K. (2018). Higher education in Cambodia: Global-local challenges and future directions. *Journal of International and Comparative Education*, 7(2), 137–152.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1), 1–17. <https://doi.org/10.61414/jeti.v5i1.103>
- Hyland, K. (2019). *English for academic purposes: An advanced resource book*. Routledge.

- Jack, H., Canavan, M., Ofori-Atta, A., Taylor, L., & Bradley, E. (2013). Recruitment and retention of mental health workers in Ghana. *PLOS ONE*, 8(2), e57940. <https://doi.org/10.1371/journal.pone.0057940>
- Jantakoon, T., Jantakun, T., Jantakun, K., Pongpanich, W., Pasmala, R., Wannapiroon, P., & Nilsook, P. (2025). The effectiveness of artificial intelligence in English instruction for speaking and listening skills: A meta-analysis. *Contemporary Educational Technology*, 17(4), ep596. <https://doi.org/10.30935/cedtech/17310>
- Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer. <https://doi.org/10.1007/978-981-4560-67-2>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon. <https://doi.org/10.1111/j.1467-971X.1982.tb00476.x>
- Lee, H., & Lee, J. H. (2024). The effects of AI-guided individualized language learning: A meta-analysis. *Language Learning & Technology*, 28(2), 134–162. <https://doi.org/10.64152/10125/73575>
- Long, D. Y., Wang, S., Md Rashid, S., & Lu, X. T. (2025). Artificial intelligence in higher education: A systematic review of its impact on student engagement and the mediating role of teaching methods. *Frontiers in Education*, 10, 1648661. <https://doi.org/10.3389/feduc.2025.1648661>
- Ma, Y., & Chen, L. (2025). Artificial intelligence in education (AIEd): Publication patterns, keywords, and research focuses. *Information*, 16(9), 725. <https://doi.org/10.3390/info16090725>
- Nadeem, M. (2025). Artificial intelligence in English language teaching: A study on vocabulary acquisition and learner engagement. *Global Journal of Educational Research and Technology*, 7(1), 112–128.
- Naidoo, K. (2017). The challenges of international students in a South African university: A case study. *Journal of Social Sciences*, 51(1–3), 51–59.
- Nguyen, H. T. (2020). Silence is gold? A study on students' silence in EFL classrooms. *International Journal of Higher Education*, 9(4), 153–167. <https://doi.org/10.5430/ijhe.v9n4p153>
- Okoli, C. I., & Nweke, K. U. (2024). Digital transformation in education: Assessing the role of AI in enhancing student research and writing. *International Journal of Digital Literacy and Digital Competence*, 15(1), 1–18.
- Oktarina, R., Sitorus, N., & Purba, D. (2024). Exploring the use of AI-driven chatbots in English for specific purposes (ESP) instruction. *Linguistics and Culture Review*, 8(S1), 312–326.
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, 14, 1255594. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Ramli, N., Mozie, N. M., Kassim, E. S., & Sari, J. D. P. (2025). Artificial intelligence in education: A bibliometric analysis of research trends and future directions. *Advances in Business Research International Journal*, 11(2), 41–55. <https://doi.org/10.24191/abrij.v11i2.8586>

- Ridha, M., Budiman, M., & Junaidi. (2026). The role of AI as a communication mediator in the learning process. *Pustaka Edukasi: Jurnal Pendidikan Indonesia*, 1(2), 73–83. <https://doi.org/10.65307/pe.v1i2.80>.
- Sari, A. W. (2025). Ethics, trust, and adoption: A literature review on student perceptions of Generative AI in Higher Education. *International Journal of Research and Applied Technology (INJURATECH)*, 5(2), 412–417.
- Seddiki, A., & Korichi, S. (2026). A theory-informed framework for selecting AI tools in language teaching. *International Journal of Technology in Education*, 9(1), 112–130. <https://doi.org/10.46328/ijte.5175>
- Shah, S., Kadir, A., & Aleksandrova, O. (2025). Beyond the prompt: Evaluating the pedagogical effectiveness of AI tutors in multicultural settings. *Interactive Learning Environments*, 33(4), 1102–1120.
- Simanjuntak, M., & Simanjuntak, J. (2025). The influence of AI-mediated communication on student-teacher rapport in online learning. *Journal of Pedagogical Research*, 9(2), 55–72.
- Tong, V. T., & Tran, H. D. (2024). AI as a scaffold: Enhancing academic writing for second language learners. *Higher Education Quarterly*, 78(4), 621–638.
- Verbeek, P. -P. (2015). Cover story: Beyond interaction: A short introduction to mediation theory. *Interactions*, 22(3), 26–31. <https://doi.org/10.1145/2751314>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wei, L., Sulaiman, T., & Ismail, N. (2024). Integrating AI in the ESL curriculum: A case study of teacher readiness and student outcomes. *Journal of Language and Linguistic Studies*, 20(1), 14–29.
- Xu, L., & Wang, Y. (2024). The impact of ChatGPT on second language learning: A mixed-methods investigation. *British Journal of Educational Technology*, 55(5), 1822–1841.
- Yi, L., & Meng, L. (2022). Teacher identity construction of non-local English-speaking EFL teachers. *Journal of Language Teaching and Research*, 13(5), 1058–1065. <https://doi.org/10.17507/jltr.1305.19>
- Yunita, W., Syafradin, S., & Haryani, H. (2025). Using AI-scaffolded practice to reduce foreign language anxiety in speaking classes. *Journal of English Language Studies*, 10(1), 12–28.
- Zhi, R., & Wang, Q. (2024). AI-supported self-regulated learning: A systematic review of the impact on academic achievement. *Journal of Educational Computing Research*, 62(4), 941–965.
- Zhu, Y., & Wang, L. (2025). Generative AI in the writing classroom: Student perceptions and performance. *Assessing Writing*, 63, 100812.