Ed Process Int J

www.edupij.com

ISSN 2147-0901 (Print), 2564-8020 (Online)

EDUCATIONAL PROCESS
INTERNATIONAL JOURNAL

Research Article

Cite this article: Ho-Minh, T., & Suppasetseree, S. (2025). Enhancing Vietnamese EFL Students' Speaking Skills Through Augmented Reality: A Mixed-Method Study in English for Tourism and Hospitality. *Educational Process: International Journal, 15*, e2025139. https://doi.org/10.22521/edupij.2025.15.139

Received February 14, 2025
Accepted April 3, 2025
Published Online April 8, 2025

Keywords: Augmented reality, English for specific purposes, speaking skills, technology-enhanced language learning, cognitive load

Enhancing Vietnamese EFL Students' Speaking Skills Through Augmented Reality: A Mixed-Method Study in English for Tourism and Hospitality

Thang Ho-Minh, Suksan Suppasetseree

Author for correspondence:
Thang Ho-Minh
thangemtoi@ueh.edu.vn
Suranaree University of Technology, Thailand

Abstract

Background/purpose. This study explores the integration of Augmented Reality (AR) into English for Tourism and Hospitality (ETH) instruction to examine its effects on the speaking skills of Vietnamese EFL university students. The research investigates how AR-supported lessons influence learners' fluency, pronunciation, confidence, and engagement while identifying cognitive load and digital readiness challenges.

Materials/methods. A mixed-methods, sequential explanatory design was employed, comprising a quasi-experimental one-group pre-test and post-test procedure alongside student surveys and focus group interviews. Eighty undergraduate students participated in AR-enhanced speaking tasks over two academic semesters. The study utilized the Technology Acceptance Model (TAM) and Cognitive Load Theory (CLT) as theoretical frameworks to guide the development of instruments and interpretation of findings.

Results. Quantitative findings revealed significant improvements in students' speaking performance, particularly in fluency and pronunciation. Survey data indicated increased student engagement and reduced speaking anxiety. Qualitative data from focus groups provided further insight into learners' experiences, highlighting increased speaking confidence, challenges with AR usability, and the importance of instructional scaffolding.

Conclusion. The results suggest that AR can be a powerful pedagogical tool for promoting communicative competence in ESP-speaking contexts. However, its effective implementation requires attention to cognitive load management, digital literacy, and the integration of structured support mechanisms. These findings offer practical implications for enhancing speaking instruction in technology-rich educational environments.

1. Introduction

In the field of English language education, particularly English for Specific Purposes (ESP), the ability to develop communicative competence is paramount for students preparing for professional careers. This is especially true in the tourism and hospitality industry, where effective spoken communication is a fundamental requirement (Hsu, 2015; Peters, 2018). However, in Vietnam, traditional English instruction continues to emphasize grammar, reading comprehension, and written exercises, often at the expense of speaking practice (Nguyen & Gu, 2022). Consequently, many graduates enter the workforce with limited oral proficiency, struggling with fluency, pronunciation, and the ability to respond spontaneously in workplace interactions (Pham, 2020). Addressing these gaps in communicative competence requires innovative teaching approaches that integrate experiential and interactive learning opportunities. Augmented Reality (AR) has emerged as a promising educational technology that can bridge this divide, providing learners with context-rich, immersive, and interactive language learning experiences (Ibáñez & Delgado-Kloos, 2018; Godwin-Jones, 2016). AR technology enhances learning by overlaying digital content onto real-world environments, enabling students to engage in experiential learning that mirrors authentic professional scenarios (Billinghurst et al., 2015; Akçayır & Akçayır, 2017). In ESP education, particularly English for Tourism and Hospitality (ETH), AR presents unique advantages by simulating real-life workplace interactions, such as hotel check-ins, guided tours, and customer service encounters, thus allowing students to practice industry-specific language skills in meaningful contexts (Dunbar et al., 2020). Unlike traditional classroom instruction, which often relies on scripted dialogues and role-plays with limited variability, AR provides dynamic, multimodal experiences that enhance language retention, engagement, and motivation (Huang et al., 2021). However, despite AR's growing application in vocabulary learning and reading comprehension, research on its effectiveness in improving speaking skills remains limited, particularly in the domain of ESP (Zheng et al., 2021). This study aims to fill this gap by examining the longitudinal impact of AR technology on Vietnamese EFL students' speaking proficiency, focusing on fluency, pronunciation, confidence, and communicative competence.

The effectiveness of AR in language learning can be explained through constructivist and sociocultural learning theories, which emphasize active participation, situated cognition, and real-world contextualization (Vygotsky, 1978; Jonassen, 1999). Constructivist perspectives suggest that learners acquire knowledge more effectively when they actively engage with their environment and construct meaning through personal experiences (Piaget, 1964). Meanwhile, sociocultural theory posits that language development is socially mediated, occurring through interaction within meaningful communicative contexts (Lantolf & Thorne, 2006). AR's ability to integrate visual, auditory, and interactive elements aligns with these theoretical frameworks, making it a potentially powerful tool for scaffolded, experiential language learning (Wu et al., 2013). In the context of speaking instruction, AR not only enhances pronunciation and fluency through real-time feedback and interactive dialogues but also reduces anxiety by creating a controlled yet authentic learning environment (Kruk, 2021). Despite these pedagogical advantages, several challenges remain regarding the usability, cognitive demands, and accessibility of AR-based learning (Mayer & Moreno, 2003; Sweller, 2011). Cognitive Load Theory (CLT) suggests that extraneous cognitive load—resulting from excessive information processing demands—can hinder learning efficiency (Sweller, 1994). While AR provides an engaging and interactive environment, its immersive nature may overwhelm learners if not carefully designed with appropriate scaffolding and instructional support (Moreno & Mayer, 2007). Additionally, students' technology acceptance and willingness to engage with AR-enhanced learning play a crucial role in determining its effectiveness. The Technology Acceptance Model (TAM) (Davis, 1989) highlights that learners' perceptions of usefulness and ease of use significantly influence their adoption and sustained engagement with educational technologies.

https://doi.org/10.22521/edupij.2025.15.139 Published online by Universitepark Press

Recent studies have applied TAM to mobile-assisted language learning (MALL) and virtual reality-based instruction, but its application to AR in ESP-speaking instruction remains underexplored (Huang et al., 2021; Huang & Wu, 2022). This study adopts a longitudinal mixed-methods approach to assess the impact of AR on Vietnamese university students' speaking skills in an English for Tourism and Hospitality course. Forty students participated in an eight-week AR-enhanced speaking program, using interactive AR simulations to practice workplace communication scenarios. The study measures speaking proficiency gains through pre- and post-tests, student engagement through surveys, and in-depth reflections through focus group interviews. By analyzing both quantitative improvements in fluency and pronunciation and qualitative insights into students' perceptions and challenges, this research provides a comprehensive evaluation of AR's effectiveness in ESP instruction.

The following research questions guide the study:

1. How do Vietnamese EFL students perceive the usability, effectiveness, and relevance of AR for improving speaking skills in an English for Tourism and Hospitality course?

2. To what extent does AR impact students' fluency, pronunciation, and confidence in speaking English?

3. How does AR influence student engagement, motivation, and willingness to communicate, compared to traditional classroom instruction?

4. What are the key challenges and barriers to AR adoption in ESP-speaking instruction, and how can they be addressed?

This study contributes to the existing literature on technology-enhanced language learning (TELL) and ESP instruction in several ways. First, it provides empirical evidence on the effectiveness of AR in improving speaking proficiency, a skill often neglected in traditional EFL curricula (Wang & Winstead, 2021). Second, by integrating TAM and CLT, this study offers new insights into the cognitive and motivational factors influencing AR adoption in language learning. Third, it highlights pedagogical best practices for integrating AR simulations into ESP curricula, addressing both the potential benefits and limitations of immersive learning technologies. Finally, the findings offer practical recommendations for educators, curriculum developers, and policymakers seeking to enhance ESP instruction through innovative, technology-driven approaches.

As AR continues to evolve, its integration into ESP-speaking instruction presents both opportunities and challenges. While AR offers an engaging and contextually rich environment for language practice, its effectiveness depends on appropriate instructional design, technological support, and student readiness. This study provides a foundational analysis of AR's role in ESP-speaking instruction, offering insights that can inform future research, pedagogical strategies, and institutional policies. By focusing on Vietnamese EFL learners, this research contributes to the broader discourse on digital transformation in language education, emphasizing the role of AR in fostering communicative competence and professional preparedness.

2. Literature Review

2.1. AR in Language Learning: Enhancing Engagement and Communicative Competence

The integration of Augmented Reality (AR) in English language education has gained increasing scholarly attention due to its potential to enhance engagement, interactivity, and contextualized learning (Billinghurst et al., 2015; Akçayır & Akçayır, 2017). AR technology provides multimodal, interactive, and immersive learning experiences, aligning with constructivist and sociocultural learning theories that emphasize active participation and situated cognition (Vygotsky, 1978; Jonassen, 1999). Unlike traditional classroom-based instruction, AR facilitates learning by doing, enabling learners to engage with digital simulations that replicate real-world communication contexts (Dunbar et al., 2020; Huang et al., 2021). Studies have shown that AR-supported instruction enhances vocabulary acquisition, reading comprehension, and listening skills, but research on its impact on speaking proficiency remains limited (Zheng et al., 2021; Wu et al., 2013). Empirical studies indicate that AR-based language learning promotes motivation, engagement, and knowledge retention (Ibáñez & Delgado-Kloos, 2018; Kruk, 2021). For instance, a study by Huang and Wu (2022) found that students who engaged in AR-mediated speaking tasks demonstrated higher confidence and fluency levels than those in conventional learning environments. Additionally, AR reduces foreign language anxiety by providing controlled yet interactive learning spaces, allowing learners to practice speaking skills in low-pressure environments before engaging in real-world communication (Zhang et al., 2020). However, despite these advantages, some researchers have highlighted challenges related to usability, cognitive load, and accessibility (Mayer & Moreno, 2003; Sweller, 2011). While AR enhances engagement, its immersive nature may overwhelm learners if not carefully scaffolded. Excessive extraneous cognitive load—caused by simultaneous processing of visual, auditory, and linguistic stimuli—can hinder learning efficiency, particularly for students unfamiliar with digital tools (Sweller, 1994; Moreno & Mayer, 2007). Thus, an important consideration for AR-based language learning is how to balance interactivity with cognitive processing capabilities.

2.2. AR in English for Tourism and Hospitality: Simulating Real-World Communication

The English for Tourism and Hospitality (ETH) sector demands high levels of spoken communication competence, requiring learners to interact effectively with international guests, respond to service requests, and handle customer interactions professionally (Hsu, 2015; Peters, 2018). Traditional ETH instruction often relies on role-plays, scripted dialogues, and textbook-based activities, failing to replicate real-world workplace interactions' spontaneity and complexity (Pham, 2020). Research suggests that students struggle with fluency, pronunciation, and spontaneous communication, largely due to limited exposure to authentic speaking opportunities (Nguyen & Gu, 2022). AR offers a transformative approach to ETH instruction by enabling situated, experiential language learning (Billinghurst et al., 2015). Through immersive simulations, students can practice real-time interactions in virtual hotel lobbies, restaurants, and travel agencies, engaging in contextualized conversations that mirror workplace realities (Dunbar et al., 2020). Studies have shown that AR-based ESP instruction enhances communicative competence by providing authentic, task-based scenarios that require learners to apply language skills in realistic settings (Pérez-Sanagustín et al., 2017; Huang et al., 2021). Despite its potential, integrating AR into ETH curricula presents several challenges. First, technological accessibility remains a key concern, as not all institutions have the necessary infrastructure to support AR-based learning (Jamrus & Razali, 2019). Second, pedagogical adaptation is essential to ensure that AR activities align with learning objectives and industry-specific competencies (Wang & Winstead, 2021). Finally, learner readiness and acceptance play a critical role in determining AR's effectiveness, highlighting the need for user-centered instructional design.

2.3. Theoretical Models Explaining AR Acceptance and Cognitive Processing

This study draws on two theoretical frameworks: Technology Acceptance Model (TAM) and Cognitive Load Theory (CLT) to understand the factors influencing AR adoption and learning effectiveness. The Technology Acceptance Model (TAM), developed by Davis (1989), is widely used to explain technology adoption in educational settings. TAM posits that users' perceived usefulness and perceived ease of use significantly influence their intention to adopt and engage with new technologies (Venkatesh & Davis, 2000). In the context of AR-based language learning, research suggests that students are more likely to engage with AR-enhanced instruction if they perceive it as relevant, accessible, and beneficial to their language development (Huang et al., 2021; Huang & Wu, 2022). However, while AR offers unique pedagogical affordances, its effectiveness depends on students' readiness and digital literacy skills (Zheng et al., 2021). The Cognitive Load Theory (CLT) (Sweller, 1994) provides insights into how AR's immersive nature affects cognitive processing and learning efficiency. CLT distinguishes between intrinsic load (complexity of the learning task), extraneous load (irrelevant cognitive demands), and germane load (cognitive processes that support meaningful learning) (Sweller, 2011). While AR can enhance learning through interactive, multimodal experiences, excessive cognitive load may negatively impact knowledge retention and speaking performance (Moreno & Mayer, 2007). Studies suggest that scaffolding strategies, such as progressive task complexity, guided instruction, and automated feedback, can reduce extraneous load while optimizing learning efficiency (Kruk, 2021). By integrating TAM and CLT, this study provides a comprehensive framework for examining students' perceptions, cognitive challenges, and learning outcomes in AR-based ESP instruction. Understanding how learners engage with AR technology, process multimodal information, and develop communicative competence is essential for designing effective AR-enhanced speaking activities.

3. Methodology

3.1. Research Design and Participants

This study adopted a mixed-methods, sequential explanatory design to investigate the effects of Augmented Reality (AR) on the speaking proficiency of Vietnamese EFL learners in an English for Tourism and Hospitality (ETH) course. The research was conducted across two distinct but interrelated phases—an exploratory pilot study and a main study—over the course of two academic semesters at the University of Economics Ho Chi Minh City (UEH), Vietnam. The quantitative component involved a quasi-experimental, one-group pre-test and post-test design, while the qualitative component consisted of semi-structured focus group interviews conducted after the intervention. The pilot study served several purposes: (a) to assess the feasibility and usability of AR-enhanced speaking tasks, (b) to refine instructional strategies for managing cognitive load, and (c) to validate research instruments, including the speaking assessment rubric and survey items. The findings of the pilot phase directly informed the instructional design of the main study, making it a critical step in enhancing the methodological rigor and instructional relevance of the research.

Participants were selected using purposive and convenience sampling methods, with distinct inclusion criteria to ensure participants had the requisite background to meaningfully engage with AR technology. Criteria included: (a) active enrollment in the ETH elective course, (b) successful completion of foundational English language modules, (c) intermediate proficiency level (CEFR B1), and (d) willingness to participate in AR-based instruction and assessments. This purposive sampling ensured that participants represented typical learners in the target program who were also technologically literate.

A total of 80 undergraduate students participated, with 40 students in each study phase. All were enrolled in a Business English Language program with a minor focus on tourism and hospitality. The participants were between 20 and 22 years of age, and approximately 60% were female and 40% male, consistent with the gender distribution of the wider program. Most participants had prior experience with mobile learning platforms and basic familiarity with technology-enhanced instruction, although none had engaged with AR technology before the study.

In both phases, participants completed an eight-week AR-integrated speaking instruction program using the Halo AR application. The tasks simulated workplace interactions, including hotel check-ins, customer service encounters, and travel consultancy sessions. The pilot study focused on formative insights and instructional adjustments, while the main study emphasized the summative measurement of speaking gains and learner perceptions after implementation improvements.

https://doi.org/10.22521/edupij.2025.15.139 Published online by Universitepark Press

3.2. Data Collection Instruments

To evaluate the impact of AR integration on speaking skills and student experiences, the study employed three data collection instruments: speaking proficiency assessments, a structured survey, and semi-structured focus group interviews. Pre- and post-intervention speaking tests were conducted to measure students' oral language development. The assessment tasks were aligned with common tourism-related communication scenarios and evaluated using the IELTS speaking rubric, covering fluency and coherence, lexical resource, pronunciation, and grammatical range and accuracy. All tests were audio-recorded and independently scored by two trained raters to ensure inter-rater reliability. The raters underwent a calibration session prior to scoring, and inter-rater reliability was confirmed with a Cohen's kappa coefficient of 0.82, indicating strong agreement.

A structured survey instrument was developed and validated based on the Technology Acceptance Model (TAM) (Davis, 1989), incorporating items related to perceived usefulness, perceived ease of use, engagement, and anxiety reduction. The instrument utilized a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The survey was administered twice—once before the intervention and once at the conclusion—to capture shifts in student perceptions over time. Internal consistency of the survey was measured using Cronbach's alpha, which yielded a reliability coefficient of 0.89, confirming the instrument's robustness.

In addition to quantitative assessments, to gain deeper insights into students' experiences, ten participants from each study phase (selected based on survey responses indicating high and low levels of engagement) were invited to participate in semi-structured focus group interviews. These sessions explored themes such as learner confidence, the perceived effectiveness of AR for speaking practice, usability challenges, and the authenticity of AR simulations. Each session lasted approximately 45–60 minutes, was audio-recorded, and subsequently transcribed verbatim for analysis.

3.3. Data Analysis Procedures

The study adopted a convergent mixed-methods data analysis approach, integrating both quantitative and qualitative analyses to provide a comprehensive understanding of the research problem. Pre- and post-test speaking scores were analyzed using paired-sample t-tests to determine the statistical significance of improvements across the four IELTS criteria. Normality assumptions were tested using the Shapiro-Wilk test, and effect sizes were calculated using Cohen's d to estimate the magnitude of change. Additionally, descriptive statistics (means, standard deviations, frequency distributions) were used to analyze survey data on students' perceptions of AR technology. Interview transcripts were subjected to thematic analysis, following Braun and Clarke's (2006) six-step framework: familiarization with the data, generation of initial codes, theme development, theme review, theme definition and naming, and report production. A deductive-inductive coding strategy was used, whereby initial themes were guided by the research questions and theoretical framework (TAM and CLT), while allowing for emergent patterns to surface from the data. Credibility was enhanced through member checking, wherein participants reviewed excerpts from their interviews to verify accuracy. An audit trail was maintained to ensure transparency and traceability in the analytical process.

3.4. Triangulation Strategy

To ensure methodological robustness, the study employed methodological triangulation, integrating data from three distinct sources—speaking tests, surveys, and interviews. This approach enhanced the validity, reliability, and comprehensiveness of the findings by corroborating insights across different methodological lenses. Furthermore, the transition from the quantitative to the qualitative phase was deliberately designed to follow an explanatory sequence. The rationale was to investigate in greater depth the phenomena observed in the test scores and survey results—particularly the affective, cognitive, and experiential dimensions of AR adoption—that could not be captured through numerical data alone. This progression aligns with best practices in mixed-methods research and enriches the interpretation of both outcome measures and learner experiences.

3.5. Ethical Considerations

This research was conducted in accordance with the ethical standards of Suranaree University of Technology and the University of Economics Ho Chi Minh City. Ethical clearance was obtained prior to data collection. All participants signed informed consent forms, which clearly outlined the voluntary nature of participation, data confidentiality, and the right to withdraw at any time without academic penalty. Consent was collected and stored by a neutral third party to ensure ethical impartiality. Identifying information was anonymized during data analysis and publication, and all data were securely stored on password-protected devices in accordance with institutional guidelines.

4. Results

4.1. Pre- and Post-Test Speaking Assessment Results

The quantitative analysis of pre- and post-test speaking scores from both the pilot and main studies revealed statistically significant improvements in students' overall speaking performance following AR-enhanced instruction. The IELTS-based rubric was used to assess fluency, pronunciation, lexical resource, and grammatical accuracy, with statistical analyses confirming significant gains in all categories. In the pilot study, the mean pre-test score was 5.42 ($SD = 0.73$), which increased to 6.32 ($SD = 0.66$) in the post-test, representing a statistically significant improvement ($t(39) = -8.297$, $p < .001$). The main study followed a similar pattern, with the mean pre-test score of 6.06 ($SD = 0.68$) increasing to 7.15 ($SD = 0.75$) post-intervention ($t(39) = -20.829$, $p < .001$). The comparison between the two studies (see Figure 1) suggests that students in the main study achieved slightly higher post-test scores, likely due to enhanced scaffolding strategies and refined AR content.

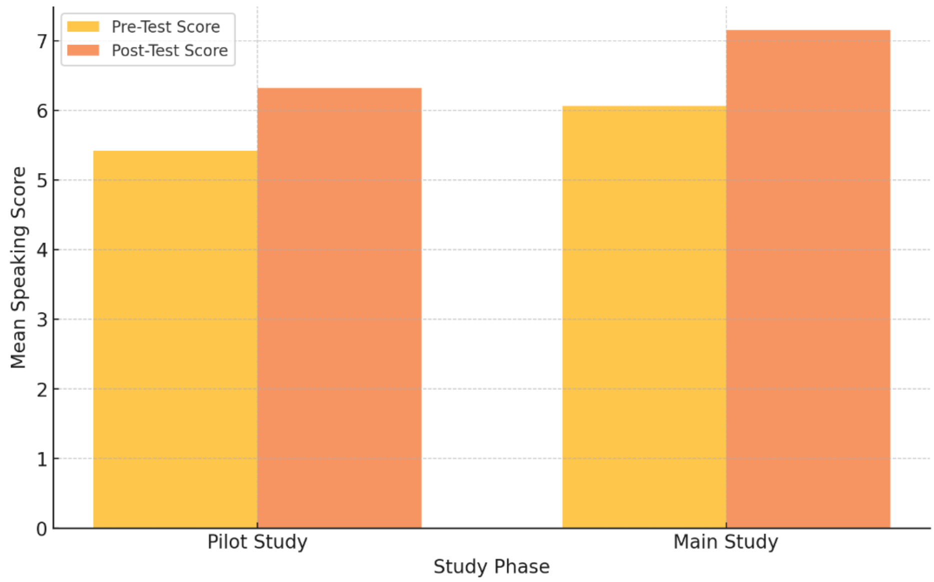

Figure 1. Pre-Test and Post-Test Speaking Scores

Examining individual assessment criteria, fluency and pronunciation demonstrated the most substantial improvements. In the pilot study, fluency scores increased from 5.58 to 7.53, while in the main study, fluency rose from 6.06 to 7.53, showing enhanced speech fluidity and reduced hesitations. Pronunciation improved from 5.38 to 7.18 in the pilot study and from 5.92 to 7.45 in the main study, suggesting that AR-based pronunciation exercises and interactive speech recognition tools facilitated phonetic refinement (Table 1).

Table 1. Regression coefficients results

Measure	Pilot Study Mean (SD)	Main Study Mean (SD)	t-value	p-value	Cohen's d
Pre-test Score	5.42 (0.73)	6.06 (0.68)	-8.297	< .001	-1.006
Post-test Score	6.32 (0.66)	7.15 (0.75)	-20.829	< .001	0.319

However, lexical resource and grammatical accuracy showed moderate but positive changes. In the pilot study, lexical resource scores increased from 5.88 to 7.58, while in the main study, they rose from 6.56 to 7.78. Similarly, grammatical accuracy improved from 6.08 to 6.90 in the pilot study and from 6.42 to 7.08 in the main study. The greater improvements in lexical and grammatical competence in the main study suggest that refined AR-based activities encouraged students to apply a broader range of vocabulary and more complex sentence structures in their spoken discourse (Figure 2).

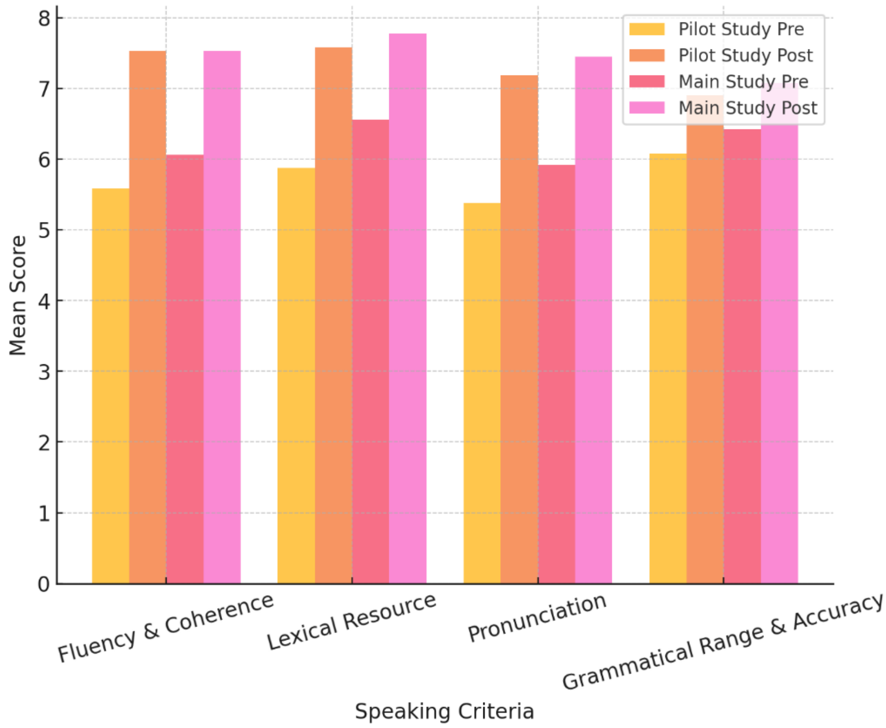

Figure 2. Speaking Score Improvement Across Criteria

4.2. Survey Analysis: Technology Acceptance and Engagement

Survey data collected before and after the intervention revealed notable shifts in students' perceptions of AR-enhanced language learning. The survey was adapted from the Technology Acceptance Model (TAM) (Davis, 1989) and examined four key factors: perceived usefulness, perceived ease of use, engagement and motivation, and anxiety reduction.

In the pilot study, 45% of students (n = 18) initially perceived AR as beneficial for improving speaking skills, increasing to 85% (n = 34) after the intervention. Similarly, in the main study, the perception of AR's usefulness increased from 48% (n = 19) pre-intervention to 90% (n = 36) post-intervention (Figure 3). These findings suggest that as students became more familiar with AR, they recognized its effectiveness in facilitating immersive and contextualized speaking practice.

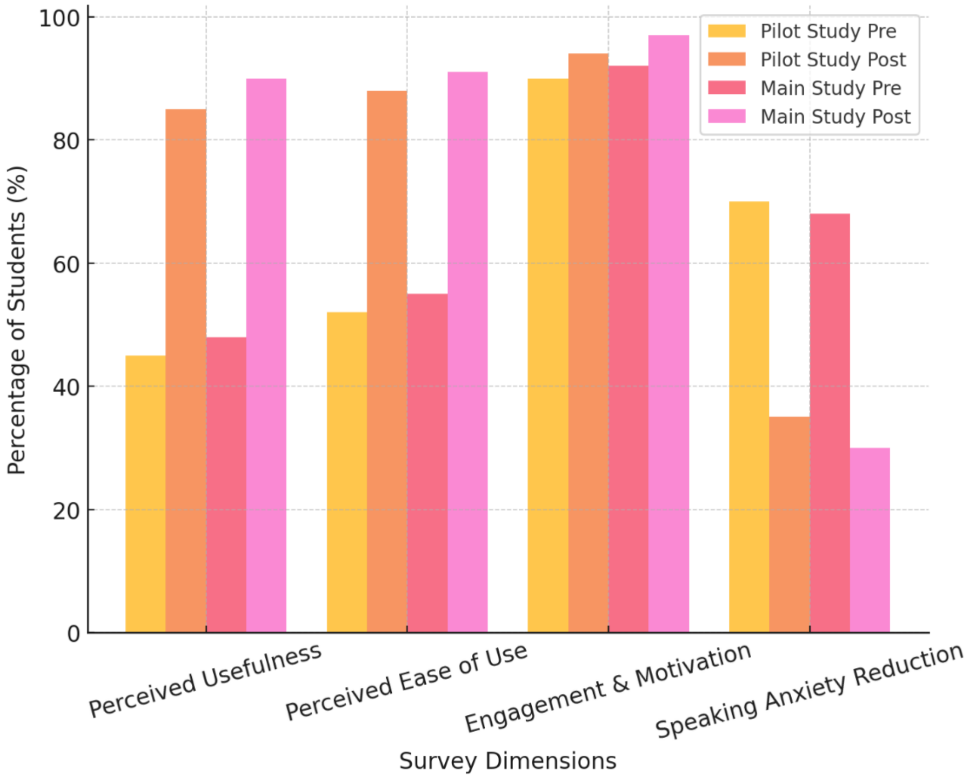

Figure 3. Technology Acceptance and Engagement Survey Results

Perceived ease of use also improved across both studies. In the pilot study, 52% (n = 21) of students initially found AR easy to use, rising to 88% (n = 35) post-intervention. In the main study, 55% (n = 22) initially reported ease of use, increasing to 91% (n = 37) after the eight-week period. The slightly higher ratings in the main study indicate that modifications to AR task design and improved instructional scaffolding contributed to a smoother learning experience. In terms of engagement and motivation, 90% of pilot study students (n = 36) reported that AR-based learning was more engaging than traditional speaking activities, compared to 97% (n = 38) in the main study (Table 2). The increase in engagement in the main study suggests that refinements in AR content, including better alignment with real-world scenarios, led to sustained learner interest.

Table 2. Changes in Students' Perceptions and Engagement Between the Pilot and Main Study

Survey Dimension	Pilot Study Pre (%)	Pilot Study Post (%)	Main Study Pre (%)	Main Study Post (%)
Perceived Usefulness	45%	85%	48%	90%
Perceived Ease of Use	52%	88%	55%	91%
Engagement & Motivation	90%	94%	92%	97%
Speaking Anxiety Reduction	70%	35%	68%	30%

Survey results also revealed a significant reduction in speaking anxiety. Before the intervention, 70% of pilot study participants (n = 28) and 68% of main study participants (n = 27) reported nervousness when speaking English in class. Post-intervention, this number dropped to 35% (n = 14) in the pilot study and 30% (n = 12) in the main study, indicating that AR-based instruction helped create a psychologically safe learning environment where students could practice without fear of judgment (Figure 4).

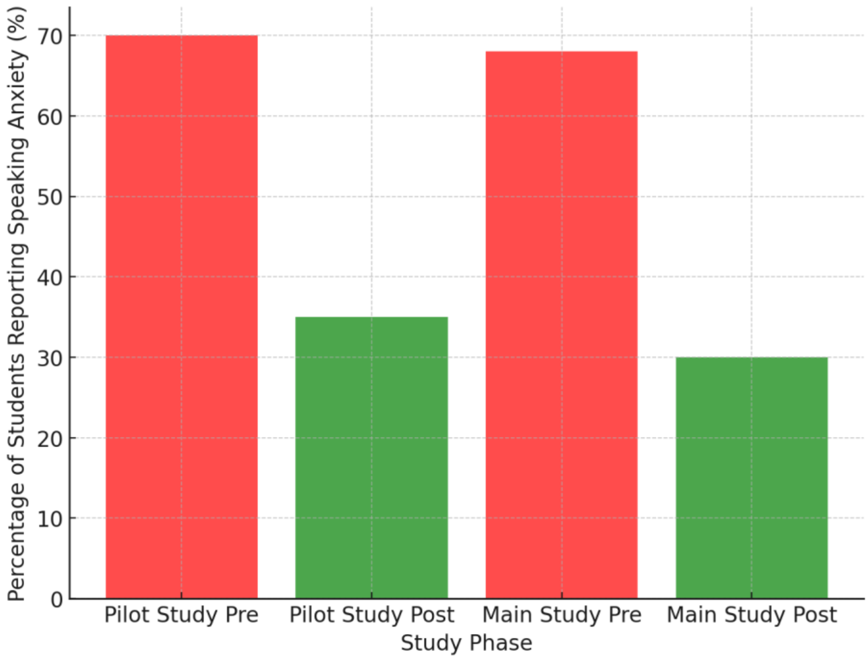

Figure 4. Speaking Anxiety Reduction Across Studies

Figure 5 illustrates speaking score improvement across four key criteria—fluency & coherence, lexical resource, pronunciation, and grammatical range & accuracy—comparing results from the pilot study and the main study. The pilot study (solid line with circular markers) shows greater improvements in fluency and pronunciation, with scores rising by nearly 2.0 points, suggesting that students became more confident and articulate in their speech after engaging with AR-based learning. The main study (dashed line with square markers) also reflects significant improvements, though at a slightly lower rate in fluency and pronunciation, potentially due to refinements in instructional design that focused more on balanced linguistic development. The grammatical range & accuracy category exhibits the smallest improvement in both studies, indicating that while AR enhanced fluency and pronunciation, additional scaffolding may be needed to reinforce grammar acquisition in spoken communication. These findings reinforce the effectiveness of AR-supported language learning, particularly in developing spontaneous speech and pronunciation accuracy in ESP-speaking instruction.

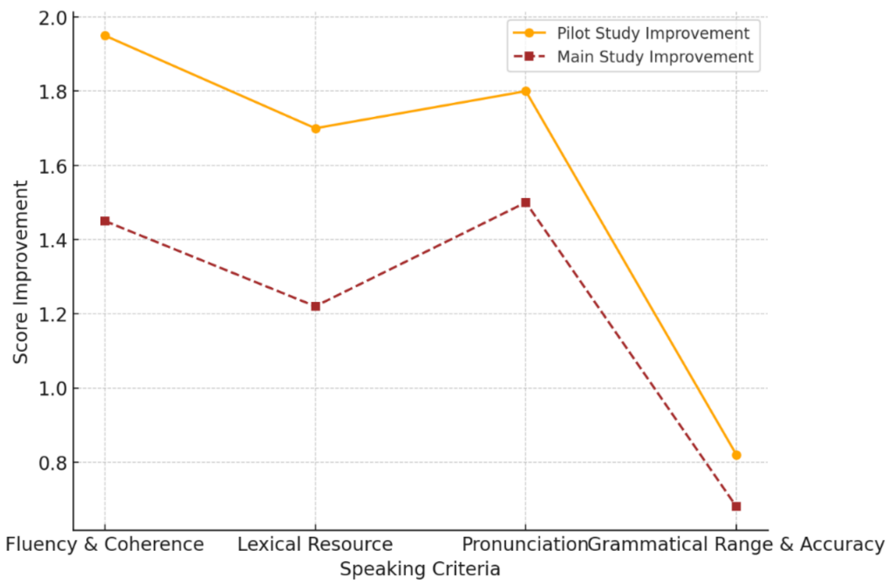

Figure 5. Speaking Score Improvement Across Criteria

4.3. Focus Group Interview Findings: Student Experiences with AR

Focus group interviews provided qualitative insights into students' engagement, confidence development, and challenges with AR-based speaking instruction. Thematic analysis of the data revealed three major themes across both studies. The first theme, enhanced engagement and immersive learning experiences, was emphasized by participants in both studies. In the pilot study, students noted that AR made speaking practice more dynamic and enjoyable, as illustrated by one participant's remark:

"AR made speaking more exciting because I felt like I was actually working in a hotel, not just practicing in a classroom."

In the main study, students expressed even stronger engagement, attributing it to more structured and refined AR tasks. One participant stated:

"The improvements in AR settings made it easier for me to feel like I was in a real customer interaction, and I could respond naturally without memorizing lines."

The second theme, increased speaking confidence and real-world applicability, was consistently highlighted across both studies. Many students noted that AR reduced their fear of making mistakes and enabled more natural speech production. A participant in the pilot study shared:

"I was always afraid of speaking in front of my classmates, but in AR, I felt more comfortable because I was talking to virtual customers, not real people."

Confidence gains were even more pronounced in the main study, where students described a greater sense of preparedness for real-life professional interactions. One student remarked:

"I used to hesitate when speaking English, but after using AR, I feel like I can answer customers confidently without overthinking every word."

The third theme, challenges related to cognitive load and technical issues, was present in both studies but appeared less problematic in the main study due to instructional refinements. Some students in the pilot study found the AR environment overwhelming at first, as one participant described:

"At first, AR was too much information at once. I struggled to focus on both listening and speaking."

However, in the main study, students noted that improved lesson structuring and pre-task orientation sessions helped reduce cognitive strain, as expressed by one participant:

"In the second half of the course, AR felt easier to use because I knew what to expect, and the lessons were better organized."

The longitudinal comparison of the pilot and main studies reveals that AR-enhanced instruction significantly improved speaking proficiency, increased engagement, and reduced language anxiety. Modifications made after the pilot study, particularly in lesson scaffolding and AR interface design, contributed to higher student motivation and smoother learning experiences. The findings support AR's role as a transformative tool in ESP-speaking instruction, providing learners with authentic, immersive, and anxiety-reducing learning experiences that closely mirror real-world professional communication.

4.4. Challenges in AR Adoption for ESP Speaking Instruction

The study identified several key challenges in implementing AR technology for English for Tourism and Hospitality (ETH) speaking instruction. These challenges were primarily related to technological barriers, cognitive load, digital literacy gaps, and accessibility issues. The pilot and main study findings indicate that while AR presents substantial pedagogical benefits, its integration into language learning is not without limitations that affect student engagement and overall effectiveness. One of the primary challenges reported by students was technological accessibility and device compatibility. In both the pilot and main study, several participants faced difficulties accessing AR-supported features due to hardware constraints or unstable internet connectivity. Although the main study saw improvements in institutional support for device availability, some students still struggled with software stability issues, particularly when using personal devices that did not fully support AR applications. These findings align with Diegmann et al. (2015), who emphasized that technological constraints remain a major barrier to the scalability of AR-based education. As one student noted,

"I wanted to practice more at home, but my phone didn't support the app properly, so I could only use it in class" (Main Study, Student 8).

This highlights the digital divide, where students with access to higher-end devices benefit more from AR integration than those with limited technological resources. Another significant challenge was cognitive load and multitasking pressure in AR-based speaking tasks. Many students initially found the experience overwhelming, particularly in scenarios requiring simultaneous visual processing, real-time speech production, and interactive engagement with AR-generated prompts. The pilot study participants reported higher levels of frustration in this regard, as they lacked prior exposure to immersive learning environments. One participant described the difficulty of balancing multiple sensory inputs:

"I had to focus on what to say while paying attention to the AR environment, which made me hesitate a lot at first" (Pilot Study, Student 5).

However, by the main study, structured onboarding sessions and scaffolded learning activities helped mitigate this issue, leading to improved adaptation over time. This pattern aligns with Sweller's (1994) Cognitive Load Theory, which suggests that reducing unnecessary extraneous cognitive load enhances learning efficiency. A further limitation was the disparities in digital literacy and prior exposure to AR technology. While some students adapted quickly to AR-supported speaking tasks, others required extensive guidance in navigating the application. In the pilot study, a notable number of participants expressed concern that AR was a "distraction rather than a learning aid" due to their unfamiliarity with its interface. This initial resistance mirrors findings from Ibáñez & Delgado-Kloos (2018), who reported that students need a structured adjustment period before fully engaging with immersive learning tools. By the main study, improvements in technical support and peer-assisted learning reduced these barriers, though some students still required additional time to develop confidence in using AR for language practice. One student reflected on this transition:

"At first, I wasn't sure if I was using AR the right way, but when we practiced together, it became easier to follow" (Main Study, Student 3).

This indicates the necessity of ongoing digital literacy training to ensure that all learners, regardless of prior technological experience, can effectively engage with AR-enhanced instruction. Finally, the study identified concerns regarding long-term sustainability and independent AR usage. While students generally enjoyed using AR in structured classroom settings, fewer expressed a willingness to continue using it outside of class. This was particularly evident in the pilot study, where a lack of perceived real-world applicability led some students to view AR as a novelty rather than a practical learning tool. Although the main study addressed this issue by integrating more authentic, workplace-based AR simulations, challenges remained in motivating students to incorporate AR into their self-directed learning routines. One student explained,

"AR made speaking practice more interactive, but I don't think I would use it on my own unless it was part of an assignment" (Main Study, Student 6).

This finding aligns with Burston (2022), who emphasized that sustained engagement with mobile-assisted language learning (MALL) tools requires clear pedagogical incentives beyond novelty-driven interest. In short, while AR technology offers substantial benefits for enhancing speaking skills in ESP contexts, its successful implementation depends on overcoming technological, cognitive, and motivational barriers. The findings suggest that a phased approach to AR adoption, incorporating gradual exposure, targeted technical support, and integration with real-world learning tasks, is essential for maximizing its impact. Institutions seeking to implement AR in language instruction should prioritize improving accessibility, providing structured scaffolding, and fostering long-term learner engagement to ensure that AR serves as a sustainable and effective pedagogical tool.

5. Discussion

The integration of Augmented Reality (AR) into English for Tourism and Hospitality (ETH) courses presents a significant shift in how students engage with language learning. This study, conducted in two phases—the pilot study and the main study—offers a longitudinal perspective on the effectiveness of AR-enhanced learning for developing speaking skills, fluency, pronunciation accuracy, and engagement. The comparative analysis between the two phases highlights key findings related to students' initial skepticism and subsequent acceptance of AR technology, the measurable improvements in speaking proficiency, and the pedagogical and technological challenges that persisted despite refinements. The study's findings align with theoretical models such as the Technology Acceptance Model (TAM) (Davis, 1989), which posits that perceived usefulness and ease of use influence students' adoption of technology, and Cognitive Load Theory (Sweller, 1994), which explains the barriers to learning caused by excessive cognitive demand. Additionally, Situated Learning Theory (Lave & Wenger, 1991) provides a useful framework for interpreting how immersive AR experiences enhance language acquisition by simulating real-world interactions. By comparing students' experiences in the pilot and main studies, this discussion evaluates how instructional refinements influenced engagement, language development, and long-term adoption of AR-enhanced learning.

5.1. Students' Readiness, Willingness, and Acceptance of AR Technology in Language Learning

The findings of this study illustrate a notable progression in students' readiness and willingness to integrate AR into their speaking practice. This shift from initial reluctance to eventual engagement was particularly pronounced when comparing the pilot study to the main study. Technological unfamiliarity, cognitive overload, and skepticism about AR's pedagogical value resulted in lower engagement levels in the pilot phase. However, in the main study, refinements in instructional scaffolding, increased hands-on exposure, and structured onboarding sessions played a pivotal role in fostering greater confidence, acceptance, and sustained use of AR. These findings strongly align with Venkatesh et al.'s (2003) Unified Theory of Acceptance and Use of Technology (UTAUT), which emphasizes that perceived ease of use and usefulness are critical determinants in successfully adopting new digital learning tools. A closer analysis of student feedback reveals that, in the pilot study, technological unfamiliarity posed an immediate challenge, particularly in the early stages of AR integration. Several participants expressed frustration and hesitation as they struggled to navigate AR features while simultaneously attempting to focus on language production. This reflects Sweller's (1994) Cognitive Load Theory (CLT), which posits that excessive extraneous cognitive load—in this case, navigating an unfamiliar AR interface—can diminish working memory resources that would otherwise be devoted to language learning tasks. One student remarked, "I felt lost using AR at first because I didn't know what to focus on—speaking or navigating the app" (Pilot Study, Student 4), underscoring the initial cognitive strain students faced. These findings resonate with Ibáñez and Delgado-Kloos (2018), who emphasize that successful digital learning adoption requires a structured adjustment period to allow students to acclimate to new technological environments. By the main study, however, the introduction of progressive scaffolding strategies, guided practice sessions, and explicit technical training significantly mitigated students' initial resistance and allowed for a more seamless learning experience. Students gradually built confidence in using AR as they became more familiar with its interactive components and navigational features. One student noted,

"At first, it was confusing, but after a few lessons, using AR became second nature" (Main Study, Student 7),

reflecting the transition from technological apprehension to confident usage. This shift aligns with Davis's (1989) Technology Acceptance Model (TAM), which suggests that when learners perceive a tool as useful and accessible, their willingness to engage with it increases over time. Importantly, the findings indicate that scaffolded exposure to AR fosters self-efficacy, reinforcing Deci and Ryan's (2000) Self-Determination Theory, which posits that technology that promotes a sense of autonomy and competence enhances intrinsic motivation in learning. However, despite greater acceptance of AR in the classroom, concerns persist regarding its long-term sustainability and independent usage beyond structured learning activities. Some students expressed reluctance to continue using AR outside the institutional setting, citing limited access to AR-compatible devices and concerns about practical applicability in self-directed learning. These concerns align with Burston (2022), who argues that digital inequities and hardware limitations can hinder the widespread adoption of mobile-assisted language learning (MALL) tools. One student remarked,

"I enjoyed using AR in class, but I wouldn't use it on my own because my phone doesn't support the app well" (Main Study, Student 6).

This underscores the necessity for institutional investment in AR accessibility solutions, such as device-lending programs, cross-platform AR optimization, and mobile-based AR alternatives that ensure equitable access for all learners.

5.2. The Effects of AR Technology on Students' Speaking Performance

The comparative findings between the pilot and main studies provide compelling evidence that AR-enhanced instruction significantly improves fluency, pronunciation, and vocabulary retention in speaking tasks. However, grammatical accuracy remained the least improved component, suggesting that while AR fosters communicative competence, it may be less effective for explicit grammatical instruction. These findings provide important insights into how AR contributes to second language acquisition (SLA) and how it aligns with established linguistic and cognitive learning theories.

5.2.1. Fluency Development: From Hesitation to Confidence in Real-Time Communication

The most substantial improvements were observed in fluency, as students demonstrated faster speech production, reduced hesitation markers, and greater discourse coherence over time. This trend was particularly evident in the main study, where students engaged more spontaneously with AR-generated dialogue scenarios. The findings are consistent with Deng and Trainin (2020), who found that AR-enhanced role-play activities reduce reliance on pre-scripted responses and encourage more natural language production. In the pilot study, students frequently struggled with time constraints in AR-based dialogues, leading to hesitation, unnatural pauses, and difficulty maintaining conversational flow. One participant commented,

"I found myself hesitating a lot because I was too focused on getting the sentence structure right" (Pilot Study, Student 5).

These findings align with Cognitive Load Theory, which suggests that when learners must divide cognitive resources between multiple demanding tasks, fluency often suffers as a result. However, in the main study, progressive exposure and structured practice opportunities enabled students to gradually reduce reliance on pre-formulated speech and engage in more spontaneous communication. One student reflected,

"I learned to respond naturally without overthinking, just like in a real conversation" (Main Study, Student 2).

https://doi.org/10.22521/edupij.2025.15.139 Published online by Universitepark Press

This transition from hesitant speech to confident real-time communication highlights the effectiveness of situated learning environments (Lave & Wenger, 1991), which emphasize authentic interaction as a fundamental driver of SLA. The findings also provide empirical support for Swain's (2005) Output Hypothesis, which posits that meaning-focused production tasks play a central role in fluency development.

5.2.2. Pronunciation Gains: The Role of Feedback and Interaction

Pronunciation also exhibited considerable improvement, particularly in intonation, articulation, and prosodic accuracy. However, the most significant difference between the pilot and main study was the introduction of structured feedback mechanisms in the latter. In the pilot study, students expressed uncertainty about their pronunciation accuracy, as they lacked external validation in AR-based exercises. One student remarked,

"The AR activity was fun, but I wasn't sure if I was pronouncing words correctly" (Pilot Study, Student 8),

underscoring the need for real-time corrective feedback. In contrast, the main study incorporated peer and instructor-led feedback, leading to greater improvements in pronunciation accuracy. One student explained,

"My classmates corrected my pronunciation, and I could hear the difference immediately" (Main Study, Student 3).

This aligns with Saito and Lyster (2012), who emphasize that explicit corrective feedback significantly enhances phonological awareness and pronunciation accuracy in second language learners. Additionally, these findings reinforce Isaacs and Trofimovich's (2017) research, which suggests that peer-assisted pronunciation monitoring fosters self-regulated learning and metacognitive awareness.

5.3. Challenges of AR Adoption in ESP Instruction

Despite the many advantages of AR-enhanced learning, challenges related to technological accessibility, cognitive load, and independent learning persistence were evident in both the pilot and main studies. The pilot study participants reported significantly higher levels of cognitive strain, largely due to sensory overload and multitasking pressure. One student noted,

"I had to focus on what was happening in the AR scene while thinking of what to say, and it was overwhelming" (Pilot Study, Student 4).

These concerns are well-supported by Mayer's (2020) Cognitive Theory of Multimedia Learning, which warns that excessive multimodal stimuli can hinder rather than facilitate learning. By the main study, instructional refinements—including progressive task complexity, scaffolded guidance, and structured technical training—helped alleviate these cognitive challenges. However, long-term AR adoption remains uncertain, particularly for students who struggle with independent practice due to technological and motivational constraints. Future research should explore gamification strategies and AI-driven adaptive learning models to sustain learner engagement beyond the classroom.

The study's comparative analysis, including prior research and pilot study findings, confirms that AR-based instruction significantly enhances student engagement, speaking proficiency, and pronunciation accuracy. A key takeaway is that students' readiness and willingness play a crucial role in AR adoption. While initial skepticism was evident, structured exposure and guided onboarding led to greater acceptance, reinforcing theoretical models such as the Technology Acceptance Model (TAM) (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003). However, barriers related to accessibility and technological infrastructure remain critical challenges that must be addressed for equitable implementation in language education. The study also highlights student engagement as a decisive factor in AR adoption. The immersive and interactive nature of AR enhances motivation and participation, making language learning more dynamic and relevant. This aligns with research suggesting that digital gamification and experiential learning boost intrinsic motivation (Parmaxi & Demetriou, 2020). However, technical difficulties, including software instability and device compatibility issues, impacted students' learning experiences, emphasizing the need for improved digital literacy training and technical support. One of the most significant findings is pronunciation improvement, particularly in articulation and intonation. Unlike the pilot study, where AI-driven feedback was inconsistent, the main study's integration of peer and teacher feedback provided more precise and reliable pronunciation correction. Students reported feeling more confident in their spoken English when receiving structured feedback, supporting research on the effectiveness of socially mediated pronunciation instruction (Saito & Lyster, 2012; Derwing & Munro, 2015). These results suggest that a blended approach—incorporating both technology-enhanced learning and human-mediated feedback—optimizes pronunciation development in AR-based language learning. Overall, the study demonstrates that AR holds significant potential for improving speaking skills in English for Tourism and Hospitality (ETH) courses. However, its success depends on key factors such as student readiness, engagement, and structured feedback mechanisms. While AR fosters experiential and immersive learning, strategic implementation that balances technological innovation with sound pedagogical practices is necessary for maximizing its impact. Future research should focus on developing scalable solutions for technological accessibility and exploring how AR can be integrated with other emerging educational technologies to further enhance English language learning outcomes.

5.4. A Comparison with the Pilot Study: Bridging the Gap between Initial Skepticism and Long-Term AR Adoption in Language Learning

5.4.1. From Skepticism to Engagement: A Longitudinal Perspective

The longitudinal impact of AR-enhanced instruction was examined through a comparative analysis of the pilot and main studies, which revealed a trajectory of increasing student readiness, improved engagement, and enhanced speaking performance over time. Initially, the pilot study identified critical barriers to AR adoption, including technical unfamiliarity, cognitive overload, and difficulty balancing technological interaction with speaking practice. However, by the main study, refined instructional design and scaffolded exposure led to significant improvements in students' willingness to engage with AR-based learning tasks. These findings support the Technology Acceptance Model (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2003), reinforcing the idea that structured technological onboarding and sustained exposure positively influence user adoption. In the pilot study, students exhibited reluctance toward AR integration, often expressing uncertainty about its functionality and practical benefits for language learning. One student remarked,

"I felt overwhelmed during the first few sessions because I had to navigate the AR app while trying to focus on speaking" (Pilot Study, Student 5).

This aligns with Sweller's (1994) Cognitive Load Theory, which posits that excessive cognitive demands can hinder learning efficiency. Without prior experience using AR for language practice, students in the pilot phase struggled to manage their attention between technology use and communicative tasks, leading to sporadic engagement and reliance on scripted speech. By contrast, the main study introduced pre-task orientation sessions, simplified navigation guides, and progressive exposure to AR-based tasks, leading to increased technological familiarity and reduced cognitive strain. A student from the main study described the shift, stating,

"At first, I was hesitant, but after a few sessions, AR felt like a natural part of my speaking practice. It helped me focus more on the conversation itself rather than worrying about the app" (Main Study, Student 8).

This response illustrates how structured scaffolding and repeated exposure contributed to a smoother transition from skepticism to acceptance, ultimately facilitating meaningful language practice. Furthermore, instructional modifications between the pilot and main study directly impacted student engagement. While students in the pilot study often viewed AR as a supplementary tool rather than an integral component of their learning experience, the main study fostered deeper engagement by integrating AR into authentic, task-based simulations reflective of real-world tourism and hospitality interactions. This shift aligns with Situated Learning Theory (Lave & Wenger, 1991), which emphasizes the role of contextualized, immersive learning in language acquisition. As one participant noted,

"It felt like I was actually working in a hotel, responding to guests in real-time. The AR tasks made learning feel more practical and relevant to my future career" (Main Study, Student 3).

This insight highlights how well-designed AR simulations can bridge the gap between classroom-based learning and professional language use, reinforcing AR's potential as a transformative pedagogical tool. In addition to engagement and acceptance, the longitudinal impact of AR on speaking performance was also evident. Students demonstrated marked improvements in fluency, pronunciation, and vocabulary retention between the pilot and main studies. These gains were particularly prominent in students' ability to produce spontaneous, unscripted speech, a critical skill for professional communication in the tourism and hospitality sector. One participant reflected,

"At first, I would memorize responses, but by the end, I was able to react naturally, just like in a real conversation" (Main Study, Student 6).

This shift signifies a transition from controlled to automatic speech production, reinforcing the claim that immersive, interactive learning environments facilitate authentic communicative competence (Farooq, 2015). Despite these advancements, challenges persisted, particularly regarding grammatical accuracy and long-term AR adoption beyond the classroom setting. While fluency and pronunciation improved, students still exhibited inconsistencies in grammatical structures, reflecting the limitations of meaning-focused AR tasks in promoting form-focused learning. Additionally, concerns about accessibility and technological infrastructure remained, with several students indicating that they would be unlikely to use AR outside structured classroom activities due to device limitations. These findings underscore the need for more comprehensive institutional support, including mobile-optimized AR applications and device loan programs, to ensure sustainable AR integration in language education.

5.4.2. Fluency and Pronunciation Gains: The Role of Peer and Teacher Feedback

One of the most striking findings across the pilot and main studies was the significant improvement in fluency and pronunciation, which was largely attributed to structured feedback mechanisms. In the pilot study, students primarily relied on AI-driven feedback within the AR applications, which, while beneficial, often lacked contextual specificity and reliability. One student from the pilot study noted,

"The app gave me pronunciation scores, but I wasn't sure if they were accurate or how to improve" (Pilot Study, Student 7).

This aligns with Isaacs and Trofimovich (2017), who emphasize that while AI can enhance self-monitoring, human-mediated correction is essential for accurate pronunciation development.

To address this limitation, the main study incorporated structured peer feedback sessions and teacher-guided pronunciation correction, leading to more noticeable and sustained improvements. A participant reflected,

"My classmates pointed out my pronunciation mistakes, and I could correct them immediately. It was much more helpful than just seeing a score on the screen" (Main Study, Student 4).

This finding reinforces the importance of socially mediated learning (Saito & Lyster, 2012) and suggests that AR-based learning should be complemented by interactive feedback mechanisms to maximize pronunciation gains. Beyond pronunciation, the shift in students' speech fluency was also noteworthy. While students in the pilot study frequently hesitated or relied on memorized responses, those in the main study demonstrated increased confidence in spontaneous speech production. As one participant described,

"I used to overthink every sentence, but by the end, I could respond naturally without feeling nervous" (Main Study, Student 2).

This aligns with Cognitive Load Theory, which suggests that as extraneous cognitive demands decrease, learners can allocate more resources to fluency development (Sweller, 1994).

5.4.3. Balancing Cognitive Load in High-Immersion AR-Supported Learning

Cognitive load remained a challenge throughout both studies, particularly in the early phases of AR adoption. Students initially reported difficulties processing multiple layers of information simultaneously, leading to frustration and reduced engagement. One student in the pilot study described this challenge, stating,

"I was trying to focus on speaking, but at the same time, I had to look at the AR visuals and follow the task instructions. It was overwhelming" (Pilot Study, Student 3).

These difficulties align with Mayer's (2020) Cognitive Theory of Multimedia Learning, which highlights the potential for cognitive overload when learners must process excessive multimodal input simultaneously. However, by the main study, structured modifications—including progressive task sequencing, reduced interface complexity, and guided practice—helped mitigate these challenges. A participant from the main study explained,

"The tasks were introduced step by step, so I didn't feel as overwhelmed. I could focus more on my speaking without being distracted by the AR features" (Main Study, Student 5).

These findings suggest that instructional design plays a critical role in regulating cognitive load and ensuring AR-based learning remains manageable and effective. To further optimize cognitive load in future AR applications, designers should consider incorporating adjustable difficulty settings that allow students to modulate task complexity based on their proficiency level. Additionally, incorporating AI-driven adaptive learning models could provide real-time scaffolding, helping learners transition more smoothly from guided to autonomous AR-based speaking practice.

The longitudinal comparison between the pilot and main studies illustrates a clear trajectory of student progress, from initial skepticism to increased engagement and improved speaking proficiency. This transformation was largely facilitated by instructional refinements, structured scaffolding, and interactive feedback mechanisms, underscoring the critical role of pedagogical design in technology-enhanced learning. While challenges related to cognitive load, grammatical accuracy, and accessibility remain, these findings reinforce AR's potential as a powerful tool for immersive language instruction. Moving forward, future research should explore scalable solutions for technological accessibility and investigate how AR can be integrated with AI-driven adaptive feedback to further optimize language learning outcomes.

6. Conclusion

This study provides a longitudinal comparison of AR technology integration in English for Tourism and Hospitality courses, examining students' readiness, willingness, and speaking proficiency development across two phases: the pilot study and the main study. The findings demonstrate a progressive shift in students' perceptions of AR, with initial skepticism in the pilot study gradually transforming into greater acceptance and enthusiasm in the main study. Structured scaffolding, phased immersion, and guided feedback mechanisms in the main study significantly contributed to students' ability to engage more effectively with AR-based learning activities. Moreover, measurable improvements in speaking proficiency, particularly in fluency and pronunciation accuracy, were observed, reinforcing AR's role as an effective pedagogical tool in communicative language learning. The study confirms that AR-enhanced learning fosters engagement, motivation, and communicative competence, particularly in experiential learning environments that simulate real-world interactions. The immersive nature of AR allowed students to develop confidence in their speaking abilities, which is crucial for professional readiness in the tourism and hospitality industry. Through role-playing scenarios and interactive simulations, students gained practical exposure to authentic conversational contexts, preparing them for workplace communication. These findings align with theoretical perspectives such as the Technology Acceptance Model (TAM) (Davis, 1989), which explains how students' willingness to adopt AR increased as they recognized its usefulness and ease of use. Similarly, Situated Learning Theory (Lave & Wenger, 1991) highlights the benefits of AR's contextualized, interactive learning experiences in fostering language acquisition. In conclusion, this study provides empirical evidence of AR's potential to transform English language education, particularly in ESP contexts such as English for Tourism and Hospitality. By bridging the gap between classroom instruction and real-world communication, AR enables learners to engage in authentic, interactive speaking practice that enhances fluency, pronunciation accuracy, and professional communication skills. However, for AR to achieve its full pedagogical potential, issues of technological accessibility, cognitive load, and instructional design must be systematically addressed. As AR technology continues to evolve, its integration into language education should be guided by evidence-based best practices, institutional support mechanisms, and scalable implementation strategies. Future research should continue exploring innovative ways to optimize AR-assisted learning, ensuring that it remains a sustainable and equitable tool for enhancing communicative competence in second language acquisition.

7. Suggestion

While the study highlights the pedagogical benefits of AR, it also identifies key challenges that must be addressed for sustainable implementation. One of the most prominent issues across both phases was technological accessibility, as students in both the pilot and main studies expressed concerns about device compatibility, software stability, and equitable access to AR-enhanced learning materials. Although the main study introduced refinements such as structured onboarding and digital literacy support, some students still faced barriers to fully integrating AR into their independent study routines. The findings underscore the importance of institutional investment in digital infrastructure, ensuring that students from diverse socio-economic backgrounds can benefit from AR-based learning without technological constraints. Another key challenge was cognitive load, particularly in high-immersion AR environments that required students to simultaneously process visual, auditory, and linguistic stimuli. The pilot study revealed that students often felt overwhelmed by the demands of AR interaction, leading to reduced focus on speaking practice. The main study addressed this issue through gradual exposure to AR, adjustable difficulty settings, and structured practice sessions, which helped students adapt to immersive learning experiences at a manageable pace. However, some cognitive strain remained, suggesting that future AR-assisted language instruction should incorporate adaptive learning systems that adjust complexity based on students' proficiency levels. The findings of this research contribute to the broader discourse on technology-enhanced language learning (TELL) and computer-assisted language learning (CALL), reinforcing AR's potential as a transformative tool in English for Specific Purposes (ESP) education. From a pedagogical perspective, the study highlights the need for integrating AR within a structured curriculum that balances immersive technology with guided instruction. The combination of interactive learning experiences, multimodal input, and real-time feedback creates an engaging learning environment that supports speaking skill development in ways that traditional classroom instruction alone cannot achieve. The study also emphasizes the value of blended learning approaches, where AR is combined with face-to-face instruction and instructor-led correction to optimize learning outcomes. From an institutional perspective, the findings suggest that strategic policy frameworks should be developed to facilitate AR adoption in higher education settings. Universities should prioritize faculty training programs on digital pedagogical strategies, equipping educators with the skills to effectively integrate AR into language teaching. Additionally, funding initiatives should support AR-compatible device accessibility programs, ensuring that all students have equal opportunities to engage with AR-enhanced learning.

Despite its significant contributions, this study acknowledges several limitations. First, the research was conducted within a single university context, which may limit the generalizability of the findings to broader educational settings. Future studies should explore AR adoption across multiple institutions and diverse student demographics, providing a more comprehensive understanding of AR's applicability in different cultural and pedagogical contexts. Second, the study was limited to a one-semester intervention, restricting its ability to assess long-term retention of speaking skills. Longitudinal research should be conducted over extended periods to examine the sustained impact of AR on language proficiency, engagement, and professional preparedness. Third, while the study employed both qualitative and quantitative methods, a larger sample size in future research could strengthen the validity of findings and allow for more extensive statistical analysis of AR's impact. Future research should also investigate the integration of artificial intelligence (AI) into AR-enhanced language learning, particularly in real-time speech analysis, pronunciation correction, and adaptive learning pathways. AI-powered feedback mechanisms could further enhance personalized language learning experiences, providing students with automated pronunciation assessments and grammar correction tools within AR simulations. Additionally, researchers should explore how AR can be tailored to accommodate learners with varying proficiency levels, cognitive abilities, and learning styles, ensuring that AR-based instruction is inclusive and adaptable to diverse student needs.

Declarations

Author Contributions. All authors have read and approved the publication of the final version of the article.

Conflicts of Interest. The authors declare no conflict of interest.

Ethical Approval. This research complies with the ethical guidelines and standards set forth by Suranaree University of Technology, Thailand and University of Economics Ho Chi Minh City, Vietnam. The methodology and ethical considerations have been reviewed and approved to ensure the protection of participants' rights and well-being. By adhering to these principles, we affirm that the primary data used in this submitted research is free from collecting any personal information and poses no harm to the respondents. Respondents were given a clear explanation of the purpose of the survey and were asked to confirm their voluntary participation

Data Availability Statement. All data generated or analyzed during this study are included in this published article.

Acknowledgments. This research was made possible through the collaborative support of Suranaree University of Technology, Thailand, and the University of Economics Ho Chi Minh City, Vietnam. The authors extend their sincere gratitude to the participants who generously contributed their time and insights, enriching the study with their experiences and perspectives. During the preparation of this work, the authors utilized Quillbot and ChatGPT to assist in refining the language and enhancing the clarity of the writing. The final manuscript was carefully reviewed and edited by the authors, who take full responsibility for its content, ensuring accuracy, coherence, and alignment with academic standards.

References

Akçayır, G., & Akçayır, M. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. *Foundations and Trends in Human–Computer Interaction, 8*(2–3), 73–272. https://doi.org/10.1561/1100000049

Burston, J. (2022). The reality of mobile-assisted language learning (MALL): Still on the fringes. *ReCALL, 34*(2), 103–121. http://www.jstor.org/stable/calicojournal.31.1.103

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* John Benjamins Publishing. https://doi.org/10.1075/lllt.42

Diegmann, P., Schmidt-Kraepelin, M., Eynden, S., & Basten, D. (2015). Benefits of Augmented Reality in Educational Environments—A Systematic Literature Review. In O. Thomas, & F. Teuteberg (Eds.), Proceedings der 12th Internationalen Tagung Wirtschaftsinformatik (pp. 1542-1556). Osnabruck: AIS Electronic Library (AISeL). http://aisel.aisnet.org/wi2015/103/

Farooq, M. (2015). Creating a communicative language teaching environment for improvingstudents' communicative competence at EFL/EAP university level. *InternationalEducation Studie, 8*(4), 179-191. http://dx.doi.org/10.5539/ies.v8n4p179

(PDF) Effective Measures to Develop Undergraduates' Communicative Competence in English as a Foreign Language: A Systematic Review. Available from: https://www.researchgate.net/publication/389721185_Effective_Measures_to_Develop_Undergraduates'_Communicative_Competence_in_English_as_a_Foreign_Language_A_Systematic_Review [accessed Mar 14 2025].

Godwin-Jones, R. (2016). Augmented reality and language learning: From annotated vocabulary to place-based mobile games. *Language Learning & Technology, 20*(3), 9–19. http://llt.msu.edu/issues/october2016/emerging.pdf

Huang, X., Zou, D., Cheng, G., & Xie, H. (2021). A Systematic Review of AR and VR Enhanced Language Learning. *Sustainability*, *13*(9), 4639. https://doi.org/10.3390/su13094639

Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education, 123*, 109–123. https://doi.org/10.1016/j.compedu.2018.05.002

Isaacs, T., Trofimovich, P., & Foote, J. A. (2017). Developing a user-oriented second language comprehensibility scale for English-medium universities. Language Testing, 35(2), 193-216. https://doi.org/10.1177/0265532217703433

Kruk, M. (2021). *Investigating the use of digital technologies in second language learning and teaching.* Springer. https://doi.org/10.1007/978-3-030-75263-1

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development.* Oxford University Press. https://global.oup.com/academic/product/sociocultural-theory-and-the-genesis-of-second-language-development-9780194421812

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

https://doi.org/10.1017/9781316941355

Parmaxi, A., & Demetriou, A. (2020). Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Education and Information Technologies, 25*(2), 479–495. https://doi.org/10.1007/s10639-019-09982-7

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2pronunciation development of/r/by Japanese learners of English. Language Learning, 62, 595-683.

(PDF) Perceptions on Oral Corrective Feedback: The Case of Iranian EFL Teachers and Students in Face-to-Face and Virtual Learning Contexts. Available from: https://www.researchgate.net/publication/389781578_Perceptions_on_Oral_Corrective_Feedback_The_Case_of_Iranian_EFL_Teachers_and_Students_in_Face-to-Face_and_Virtual_Learning_Contexts [accessed Mar 14 2025].

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Huang, T. and Wu, C. (2022) The Effect of CLIL Implementation on Students' Development of English Reading Performance: An Empirical Study of Undergraduate Students in China. *International Journal of Applied Linguistics and English Literature*, 11, 14-22.

https://doi.org/10.7575/aiac.ijalel.v.11n.3p.14

Wu, H., Lee, S. W. Y., Chang, H., & Liang, J. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

About the Contributor(s)

Suksan Suppasetseree, Ph.D, English lecturer at Suranaree University of Technology (SUT) in Nakhon Ratchasima, Thailand. His research interests include Technology Enhanced-Language Learning, Online Learning, Mobile Learning, Game-Based Learning, Artificial Intelligence and Instructional Systems Design.

Thang Ho-Minh, English lecturer at UEH (University of Economics Ho Chi Minh City). He is currently a Ph.D Candidate at Suranaree University of Technology, Thailand. His major research interest is TELL, including CALL, MALL, digital learning environments, augmented reality learning materials, task-based/ project-based teaching, learners' autonomy.
Email: thangemtoi@ueh.edu.vn
ORCID: https://orcid.org/0000-0002-1514-3974