



Advancing global intelligibility and intercultural competence: A gamified, Self-Determination Theory-grounded pronunciation framework for tertiary EFL learners

Supriatna Ade Wijaya¹, Farah binti Abdul Rahman¹, Gede Putu Arimbawa¹, Tran Thi Duong²

¹ Department of Education Management, Ahmad Dahlan University, Yogyakarta, Indonesia

² Department of Foreign Languages, Thai Nguyen University of Technology, Vietnam

DOI: <https://doi.org/10.66856/ijerd.2026.8.2.8033>

Abstract

This quasi-experimental study investigated the efficacy of an 8-week, gamified pronunciation instructional framework grounded in Self-Determination Theory (SDT). Utilizing an explanatory sequential mixed-methods design, the study examined changes in English intelligibility, comprehensibility, and Intercultural Communicative Competence (ICC) among 116 non-English major university students in Vietnam, divided into an experimental group and a control group. Quantitative analyses via ANCOVA revealed that the experimental cohort achieved significantly higher post-test speech intelligibility and comprehensibility scores, alongside profound shifts toward global linguistic empathy regarding World Englishes. Furthermore, psychometric scales indicated substantial drops in classroom boredom and oral communication anxiety. Qualitative thematic analysis of semi-structured interviews corroborated these outcomes, demonstrating that quest-based tasks and collaborative team scenarios successfully satisfied learners' basic psychological needs for autonomy, competence, and relatedness. This research provides a systematically validated pedagogical blueprint for seamlessly integrating linguistic development with cross-cultural pragmatic awareness within modern language learning curricula.

Keywords: Gamification, Self-Determination theory, pronunciation pedagogy, speech intelligibility, intercultural communicative competence, World Englishes

Introduction

In an increasingly globalized world, English has firmly established itself as a Global Lingua Franca (ELF), shifting the focus of pronunciation pedagogy away from the nativeness principle which once pressured learners to eradicate their accents in pursuit of an idealized native-speaker standard toward the intelligibility and comprehensibility principles (Bøhn & Hansen, 2017; Jeong & Lindemann, 2025; Levis, 2005) [3, 11, 13]. Modern scholarship argues that forced assimilation to native norms is both unnecessary and counterproductive, advocating instead for a focus on core phonological features that safeguard communicative efficacy in diverse, multilingual settings (Bouchhioua, 2017; Chen, 2016; Galante & Piccardo, 2022; Levis, 2011) [4, 6, 8, 14].

Achieving this mutual intelligibility requires modern instruction to synthesize linguistic accuracy with Intercultural Communicative Competence (ICC), ensuring learners not only master segmentals and suprasegmentals but also develop an empathetic awareness of diverse World Englishes (Bouchhioua, 2017; Chen, 2016; Luu, 2026; Luu & Nguyen, 2026) [4, 6, 15, 19]. To address the difficulty of integrating these elements, Mobile-Assisted Language Learning (MALL) and gamification have emerged as innovative methodologies. Grounded in Self-Determination Theory (SDT), gamified architectures leveraging points, badges, and challenges restructure the psychological climate of the classroom to sustain engagement and foster learner autonomy, offering a powerful mechanism to revitalize a historically neglected area of pedagogy (Barcomb & Cardoso, 2020; Luu *et al.*, 2025; Luu, Nguyen, Do, & Nguyen, 2025; Sholekhah & Fakhurriana, 2023; Shroff *et al.*, 2020; Tejedor-García *et al.*, 2020) [2, 16, 25, 27].

Despite this potential, pronunciation remains marginalized within EFL curricula, often overshadowed by grammar and vocabulary or traditional, teacher-centered drilling that triggers academic boredom and communication anxiety (Agustina *et al.*, 2025; Bouchhioua, 2017; Khelef & Khelef, 2024; Luu, Nguyen, & Nguyen, 2025^[1, 4, 12, 16]; Saito, 2013). Vietnamese university students, in particular, face persistent segmental hurdles and lack the communicative confidence necessary for real-world interaction (Luu & Le, 2026; Luu & Nguyen, 2026) [15, 19]. More critically, traditional classrooms operate in a socio-cultural vacuum, treating phonology as a purely mechanical exercise divorced from the intercultural realities of post-pandemic digital and physical spaces (Luu, 2025; Luu *et al.*, 2025) [20].

Without exposure to phonetic variations, learners fail to develop the cultural flexibility required to decode World Englishes, often defaulting to rigid stereotypes (Luu, 2026) [15]. Furthermore, while AI-driven tools offer flexible practice, existing digital applications are often designed for isolated, individualistic exercises that neglect collaborative social interaction and fail to address technical inaccuracies when processing diverse L2 accents (Agustina *et al.*, 2025; Luu, Nguyen, Do, & Nguyen, 2025) [1, 16]. This fragmented landscape leaves a distinct scholarly void regarding how a theory-driven, gamified pedagogical framework can simultaneously enhance phonetic intelligibility and intercultural communicative competence.

This study directly addresses this gap by implementing an SDT-grounded, gamified pronunciation framework designed to improve EFL learners' intelligibility while reshaping their intercultural attitudes toward diverse accents. By shifting the pedagogical focus from native-like perfection to global intelligibility, this research advances current SLA literature and provides a systematically validated blueprint for

educators to integrate collaborative, authentic intercultural scenarios into regular coursework (Agustina *et al.*, 2025; Jeong & Lindemann, 2025; Khelef & Khelef, 2024; Luu *et al.*, 2025; Luu & Le, 2026; Luu, Nguyen, Do, & Nguyen, 2025) [1, 11, 12, 15, 16]. To systematically evaluate the effectiveness of this framework, this investigation is guided by the following research questions:

1. To what extent does the integration of gamification in pronunciation instruction improve EFL learners' English intelligibility and comprehensibility?
2. How does gamified pronunciation learning affect EFL learners' Intercultural Communicative Competence, specifically regarding their attitudes towards World Englishes and English as a Lingua Franca?
3. What are the learners' perceptions and experiences regarding the use of gamified elements in developing both their pronunciation and intercultural awareness?

Methodology

Research design & context

This study addresses its research questions through a quasi-experimental, explanatory sequential mixed-methods design. A mixed-methods approach is selected because it enables a comprehensive evaluation of the pedagogical intervention, combining the statistical power of quantitative metrics with the rich contextual depth of qualitative narratives. The quantitative phase primarily utilizes a non-equivalent control group pre-test/post-test architecture to measure changes in English intelligibility, comprehensibility, and Intercultural Communicative Competence (ICC). Following the collection of quantitative data, the qualitative phase employs semi-structured interviews to capture nuanced descriptions regarding the learners' psychological shifts, academic emotions such as boredom and anxiety, and subjective experiences during the intervention.

The study was conducted over an intensive 8-week period during the second semester of the 2025^[20]–2026 academic year at a major public university located in southern Vietnam. This specific instructional context is highly representative of contemporary Vietnamese tertiary English as a Foreign Language (EFL) environments. In these settings, pronunciation instruction has historically been marginalized or confined to rigid, native-centric mechanical drilling. Consequently, students primarily interact in mono-cultural classrooms that offer minimal exposure to international variations, creating a distinct pedagogical gap despite the expanding demands of global digital and physical networks.

Participants & sampling technique

A convenience sampling technique was employed to select intact classes of non-English major university students who were enrolled in a compulsory General English course. A total of 116 students participated in the study, all of whom were pre-screened using an institutional proficiency placement exam to ensure baseline homogeneity. Their English proficiency corresponded to the B1 level of the Common European Framework of Reference for Languages (CEFR). The participants were distributed into two intact groups: the Experimental Group (EG), comprising 58 students who received the SDT-grounded, gamified pronunciation instruction, and the Control Group (CG), consisting of 58 students who underwent traditional, teacher-centered pronunciation instruction covering identical phonological features.

The comprehensive demographic profile of the participants encompassing gender, age distribution, academic major, prior exposure to gamified learning tools, and self-reported baseline pronunciation confidence is detailed in Table 1.

Table 1: Participant demographic profile

Demographic Variable	Category	Experimental Group (n=58)	Control Group (n=58)	Total Sample (N=116)
Gender	Male	22 (37.9%)	25 (43.1%)	47 (40.5%)
	Female	36 (62.1%)	33 (56.9%)	69 (59.5%)
Age	18–19 years	41 (70.7%)	43 (74.1%)	84 (72.4%)
	20–21 years	17 (29.3%)	15 (25.9%)	32 (27.6%)
Academic Major	Information Technology	20 (34.5%)	18 (31.0%)	38 (32.8%)
	Business Administration	23 (39.7%)	22 (37.9%)	45 (38.8%)
	Engineering / Tourism	15 (25.8%)	18 (31.1%)	33 (28.4%)
Prior Gamified Learning	Never used	39 (67.2%)	42 (72.4%)	81 (69.8%)
	Occasionally used	19 (32.8%)	16 (27.6%)	35 (30.2%)
Pronunciation Confidence (Self-Reported)	Low	34 (58.6%)	31 (53.4%)	65 (56.0%)
	Moderate	21 (36.2%)	23 (39.7%)	44 (37.9%)
	High	3 (5.2%)	4 (6.9%)	7 (6.1%)

Pedagogical intervention

Both groups covered the exact same pronunciation curriculum over the 8-week intervention period. The instructional contents specifically targeted core segmental features, including vowel contrasts, final consonant clusters, and pronunciation-spelling inconsistencies, alongside critical suprasegmental elements such as lexical stress, sentence stress, and linking. These features were selected because they directly determine speech intelligibility within an English as a Lingua Franca (ELF) paradigm. However, the method of pedagogical delivery differed fundamentally between the two groups.

Experimental group (gamified, sdt-grounded framework)

The pedagogical intervention for the Experimental Group was systematically engineered to support the three basic psychological needs outlined by Self-Determination Theory: autonomy, competence, and relatedness. Instruction was operationalized through a blended platform via a specialized web and mobile application interface.

To cultivate autonomous learning and competence, structural phonological rules were discovered through active digital Quests rather than passive lecturing. Students worked in small, collaborative teams termed Guilds to solve

interactive puzzles, identify acoustic traps, and bridge pronunciation-spelling ambiguities, thereby satisfying their need for relatedness. Furthermore, intercultural scenario role-plays were embedded into the coursework to simulate realistic cross-cultural misunderstandings caused by phonological errors, such as a mispronounced final consonant leading to critical lexical ambiguity in an international business meeting. Students had to actively navigate, diagnose, and resolve these communicative breakdowns, which shifted their focus away from native-like perfection toward functional global intelligibility. Immediate feedback loops were maintained through automated scoring systems where students accumulated experience points, earned badge milestones (e.g., Lingua Franca Master, Stress Navigator), and viewed team-based leaderboards designed to lower communication anxiety and alleviate academic boredom.

Control group (traditional framework)

The Control Group received conventional, instructor-led pronunciation instruction. The pedagogical sequence strictly followed a traditional Presentation, Practice, and Production (PPP) model. The instructor first introduced the target phonological features using isolated audio models reflecting native-speaker accents, specifically Received Pronunciation or General American. Students then practiced these features through mechanical drills, including choral repetition, individual drilling, and structured read-aloud tasks. The practice exercises were entirely decontextualized, utilizing word lists, static sentences, and reading passages completely

divorced from intercultural or real-world communicative contexts. Feedback in this group was delivered manually and delayed by the instructor, focusing exclusively on phonetic accuracy relative to native-speaker standards. No gamified digital rewards, adaptive progression pathways, or collaborative peer tasks were introduced during their sessions.

Data collection instruments

To comprehensively address the research questions, data were systematically gathered using five distinct instruments. In accordance with rigorous academic standards, each instrument was either adapted from an established psychometric framework or structured based on validated second language acquisition rubrics.

Speech Intelligibility and Comprehensibility Test (SICT)

The Speech Intelligibility and Comprehensibility Test (SICT) was implemented to collect oral data from participants before and after the 8-week intervention. Speech samples were recorded individually using controlled paragraph-reading tasks and spontaneous picture-description tasks. These audio samples were double-blind evaluated by three expert international raters using 9-point Likert scales adapted directly from the standard second language acquisition metrics of Munro and Derwing (1995) [22]. Intelligibility was operationalized as the percentage of words correctly transcribed by the raters, whereas comprehensibility measured the raters' perceived ease of understanding the speaker's utterance. The specific metrics of this assessment tool are summarized in Table 2.

Table 2: Specification Matrix for the Speech Intelligibility and Comprehensibility Test (SICT)

Component	Task Type	Evaluation Metric	Rater Configuration	Target Objective
Section 1	Controlled Paragraph Reading	9-point Likert Scale (Munro & Derwing, 1995) [22]	Three independent international raters	Measures actual phonetic clarity and segmentals (RQ1)
Section 2	Spontaneous Picture Description	Word Transcription Percentage & Ease of Understanding	Double-blind scoring protocol	Assesses global comprehensibility under cognitive load (RQ1)

Intercultural Communicative Competence and Accent Attitudes Scale (ICCAAS)

Socio-cultural and ideological shifts regarding language variety were measured using the Intercultural Communicative Competence and Accent Attitudes Scale (ICCAAS). This instrument consists of an 18-item questionnaire rated on a 5-point Likert scale, ranging from

Strongly Disagree to Strongly Agree. The underlying constructs were adapted from the English as a Lingua Franca framework pioneered by Jenkins (2007) [10] and the non-native accent attitude inventory. The questionnaire explicitly traces whether learners transition away from native-centric linguistic biases toward intercultural empathy and communicative flexibility, as detailed in Table 3.

Table 3: Item Distribution for the Intercultural Communicative Competence and Accent Attitudes Scale (ICCAAS)

Subscale Dimension	Core Theoretical Origin	Sample Item Theme	Number of Items
Linguistic Awareness	Jenkins (2007) [10]	Acceptability of non-native variations in global spaces	6 Items
Intercultural Empathy	Jenkins (2007) [10]	Openness to communicating with diverse English speakers	6 Items
Pedagogical Shift	Jenkins (2007) [10]	Perceived value of global intelligibility vs. native norms	6 Items

Foreign Language Classroom Anxiety Scale Pronunciation Subscale (FLCAS-P)

Psychological anxiety tied specifically to oral performance was evaluated via the Foreign Language Classroom Anxiety Scale - Pronunciation Subscale (FLCAS-P). Modified from the classic foundational anxiety scale established by Horwitz,

Horwitz, and Cope (1986) [9], this instrument isolates affective distress arising during pronunciation tasks. It features 12 items measured on a 5-point Likert scale, capturing dimensions such as communication apprehension, speech judgment anxiety, and the fear of negative evaluation by peers or instructors within the classroom environment, as outlined in Table 4.

Table 4: Dimension Matrix for the Foreign Language Classroom Anxiety Scale - Pronunciation Subscale (FLCAS-P)

Construct Dimension	Source Foundation	Measurement Focus	Number of Items
Speech Apprehension	Horwitz <i>et al.</i> (1986) ^[9]	Nervousness during oral production and reading aloud	4 Items
Social Evaluation Fear	Horwitz <i>et al.</i> (1986) ^[9]	Anxiety regarding peer judgment of pronunciation errors	4 Items
Performance Anxiety	Horwitz <i>et al.</i> (1986) ^[9]	Feelings of inadequacy compared to other English speakers	4 Items

Academic Boredom Scale in Pronunciation (ABSP)

To quantify negative achievement emotions during instruction, the Academic Boredom Scale in Pronunciation (ABSP) was administered. This instrument was structured directly upon the Control-Value Theory of Achievement Emotions formulated by Pekrun (2006) ^[23]. Comprising 10

items evaluated on a 5-point Likert scale, this psychometric tool measures situational boredom, cognitive under-stimulation, and mental disengagement induced by different instructional methods, distinguishing between traditional mechanical drilling and interactive gamified pathways as shown in Table 5.

Table 5: Measurement Domains for the Academic Boredom Scale in Pronunciation (ABSP)

Measurement Domain	Theoretical Framework	Psychological Indicator	Number of Items
Valence & Value	Pekrun (2006) ^[23]	Perceived meaninglessness or lack of purpose in tasks	3 Items
Arousal Levels	Pekrun (2006) ^[23]	Sluggishness, clock-watching, and mental fatigue	4 Items
Cognitive Avoidance	Pekrun (2006) ^[23]	Tendency to mentally disengage or distract oneself	3 Items

Semi-Structured Interview Guide

Qualitative data were gathered via a Semi-Structured Interview Guide designed to explore the psychological mechanisms underlying student experiences. The guide was formulated based on the Self-Determination Theory constructs of Deci and Ryan (2000) ^[7]. It features 6 core open.

Ended prompts alongside flexible qualitative follow-up probes. The structure is designed to elicit detailed descriptions of how specific game design mechanics supported or thwarted the students' basic psychological needs for autonomy, competence, and relatedness, while tracking their evolving perceptions of global accents, as detailed in Table 6.

Table 6: Structural Layout of the Semi-Structured Interview Guide

Core Inquiry Module	Theoretical Foundation	Target Qualitative Construct	Primary Prompt Focus
Module 1: Need Support	Deci & Ryan (2000) ^[7]	Autonomy, Competence, Relatedness	Perceptions of quest choices, leaderboards, and team guild tasks
Module 2: Emotion Shifts	Pekrun (2006) / Horwitz <i>et al.</i> (1986) ^[9, 23]	Anxiety mitigation & Boredom reduction	Subjective feelings during game loops versus previous drills
Module 3: ICC Growth	Jenkins (2007) ^[10]	Accent acceptance & Empathy development	Real-world readiness derived from cultural role-play scenarios

Data analysis procedures

The quantitative and qualitative datasets were processed independently using separate analytic software before being synthesized during the final discussion stage.

Statistical computations were conducted utilizing SPSS (Version 28.0) and AMOS. First, descriptive statistics, including Means (M) and Standard Deviations (SD), were calculated for both groups across all pre-test and post-test scores. Second, inferential statistical analyses were performed to evaluate the intervention's impact. Paired-samples t-tests were conducted to evaluate within-group improvements from pre- to post-test. To compare post-intervention performance between the Experimental Group and the Control Group, an Analysis of Covariance (ANCOVA) was executed, utilizing the baseline pre-test scores as covariates to effectively eliminate initial variations. Finally, the practical magnitude of the pedagogical framework was determined by calculating effect sizes using Cohen's d and partial eta-squared (η_p^2), with statistical significance set at $p < .05$.

Audio recordings from the semi-structured interviews were transcribed verbatim and subjected to thematic analysis, following the iterative six-phase framework established by Braun and Clarke (2006) ^[5]. The primary researcher first read the transcripts repeatedly to ensure complete immersion in the data. Initial open coding was then

conducted line-by-line to extract meaningful statements regarding specific game elements, emotional variations, and evolving socio-cultural attitudes.

In the subsequent stages, these codes were organized into broader sub-themes and macro-themes. These categories were structurally mapped onto the pre-defined theoretical dimensions of Self-Determination Theory (autonomy, competence, relatedness) and the core components of Intercultural Communicative Competence, ensuring a robust, theory-driven interpretation of the qualitative narratives.

Ethical Considerations

The design and execution of this study strictly conformed to established institutional ethical protocols for research involving human subjects. To ensure complete transparency and participant safety, a rigorous ethical protection protocol was enforced throughout the study. Prior to commencing the 8-week program, all 116 potential participants received a comprehensive written information brief explaining the specific objectives, operational scope, duration, and data use policy of the research. Signed informed consent forms were securely obtained from every participant before any data collection took place.

Participants were explicitly informed that their enrollment was entirely voluntary. They maintained the right to

withdraw from the research at any juncture during the 8-week cycle without facing any academic penalties, negative institutional evaluations, or disruptions to their regular course grading. To fully protect individual privacy, all collected data, speech recordings, survey responses, and interview transcripts were completely anonymized using a structured alphanumeric coding system (e.g., Student EG-01, Student CG-14). No personally identifiable information was attached to the analytical datasets. All files were stored on password-protected cloud storage systems accessible exclusively to the primary research team. To ensure that the students in the Control Group were not academically disadvantaged by the exclusion of the innovative methodology, all gamified learning modules, interactive role-play scenarios, and collaborative digital quest assets were made fully accessible to them via an open-access institutional repository immediately following the completion of the post-test data collection phase.

Findings and Discussions

Quantitative and Qualitative Results for Research Question 1

Research Question 1 sought to evaluate the extent to which the integration of gamification in pronunciation instruction improves EFL learners' English intelligibility and

comprehensibility. To establish baseline equivalence, independent samples *t*-tests were performed on the pre-test scores of the Speech Intelligibility and Comprehensibility Test (SICT). The results revealed no statistically significant differences between the Experimental Group (EG) and the Control Group (CG) for either intelligibility ($t(114) = 0.42$, $p = .675$) or comprehensibility ($t(114) = 0.51$, $p = .611$), confirming that both groups possessed comparable oral production skills prior to the pedagogical intervention.

Following the 8-week instructional period, a one-way Analysis of Covariance (ANCOVA) was conducted to compare the post-test scores between the two cohorts while statistically controlling for pre-test performance. The inferential analysis demonstrated a highly significant main effect for the instructional treatment on post-test speech intelligibility ($F(1, 113) = 42.61$, $p < .001$, $\eta^2 = .274$), indicating that the gamified framework explained 27.4% of the variance in the learners' phonetic clarity. Similarly, the ANCOVA for speech comprehensibility yielded a robust, statistically significant superiority for the EG over the CG ($F(1, 113) = 51.18$, $p < .001$, $\eta^2 = .312$). Paired samples *t*-tests further revealed that while the CG exhibited nominal, non-significant improvements, the EG achieved significant gains from pre-test to post-test in both domains. The complete statistical breakdown is presented in Table 7.

Table 7: ANCOVA and Descriptive Results for Speech Intelligibility and Comprehensibility

Variable	Group	Pre-test M(SD)	Post-test M(SD)	Adj. Mean	F-value	p-value	Partial η^2
Intelligibility	Experimental (n=58)	4.82 (0.65)	6.74 (0.58)	6.72	42.61	< .001	.274
	Control (n=58)	4.77 (0.71)	5.21 (0.64)	5.23			
Comprehensibility	Experimental (n=58)	4.51 (0.59)	7.12 (0.61)	7.15	51.18	< .001	.312
	Control (n=58)	4.46 (0.63)	5.04 (0.55)	5.01			

The qualitative interview data corroborated these statistical outcomes, with participants noting that the instant feedback loops provided by the automated scoring engine explicitly targeted segmental distortions and structural word-stress ambiguities that had long hindered their vocal clarity.

"In previous English classes, I was always afraid to speak because I never knew if my final consonants or word stress were correct. The automated quest scoring system pointed out my errors immediately without embarrassing me. Practicing to beat the system's score made me realize how mispronouncing final sounds completely changed the words I was trying to say." (Participant EG-12)

"Working through the digital challenges made me actively focus on word linking and sentence stress. I used to read English like Vietnamese, very flat and word-by-word. The app forced me to look at the rhythm of the sentence to advance to the next level, and the raters' post-test scores showed that I am much easier to understand now." (Participant EG-44)

Quantitative and Qualitative Results for Research Question 2

Research Question 2 examined how the gamified pronunciation framework altered EFL learners' Intercultural Communicative Competence (ICC), focusing specifically on their ideological attitudes toward World Englishes and English as a Lingua Franca (ELF). The quantitative shifts were measured through the 18-item Intercultural Communicative Competence and Accent Attitudes Scale (ICCAAS). An ANCOVA, with pre-test scores entered as the covariate, verified a profound statistical divergence between the two groups by the end of the 8 weeks ($F(1, 113) = 84.15$, $p < .001$, $\eta^2 = .427$).

The experimental cohort manifested an exceptional shift from native-centric evaluative biases toward a paradigm of global intercultural empathy. Conversely, the control group's attitudes remained largely stagnant, continuing to view non-native phonetic features through a deficit lens. The pre- and post-test values across the three core dimensions of the ICCAAS are detailed in Table 8.

Table 8: Descriptive and Inferential Statistics for the ICCAAS Dimensions

ICCAAS Dimension	Group	Pre-test M(SD)	Post-test M(SD)	Mean Difference	F-value	p-value
Linguistic Awareness	Experimental	2.45 (0.55)	4.21 (0.48)	+1.76	72.33	< .001
	Control	2.51 (0.51)	2.68 (0.53)	+0.17		
Intercultural Empathy	Experimental	2.31 (0.62)	4.38 (0.51)	+2.07	91.04	< .001
	Control	2.38 (0.58)	2.52 (0.60)	+0.14		
Pedagogical Shift	Experimental	2.18 (0.49)	4.15 (0.54)	+1.97	68.50	< .001
	Control	2.22 (0.52)	2.35 (0.47)	+0.13		

The qualitative transcriptions illuminated the cognitive mechanisms behind these numbers. Students explained that navigating the embedded intercultural scenario role-plays dismantled their rigid preconceptions regarding the supremacy of Inner-Circle accents, reframing diverse World Englishes as valid, functionally legitimate communication tools.

"Before this course, I genuinely believed that any accent that did not sound perfectly American or British was 'bad English'. The cultural role-play simulations where we had to resolve business misunderstandings caused by different regional accents made me see that intelligibility is what matters most. I learned to listen to the speaker's meaning rather than judging their background." (Participant EG-07)

"The team-based simulation tasks forced us to adapt to various English accents from India, Singapore, and Europe. It made me realize that English belongs to everyone who uses it globally. I no longer feel ashamed of my Vietnamese accent as long as my pronunciation is clear, structured, and mutually intelligible." (Participant EG-31)

Quantitative and qualitative results for research question 3

Research Question 3 investigated learners' perceptions and experiences regarding the use of these gamified elements in developing both their pronunciation and intercultural awareness, focusing specifically on their internal psychological needs and emotional adjustments. Quantitative variations in academic emotions were monitored using the Foreign Language Classroom Anxiety Scale - Pronunciation Subscale (FLCAS-P) and the Academic Boredom Scale in Pronunciation (ABSP).

The independent ANCOVA computations revealed that the experimental group experienced a significant drop in both performance anxiety ($F(1, 113) = 63.44, p < .001, \eta^2 = .359$) and classroom boredom ($F(1, 113) = 92.11, p < .001, \eta^2 = .449$) relative to the control group. This confirms that the gamified environment generated a highly protective, emotionally supportive atmosphere for language production. These emotional outcomes are summarized in Table 9.

Table 9: Post-Intervention Psychological and Emotional Measures

Psychometric Scale	Group	Pre-test M(SD)	Post-test M(SD)	Mean Change	F-value	p-value	Partial η^2
Anxiety (FLCAS-P)	Experimental	3.92 (0.54)	2.15 (0.46)	-1.77	63.44	< .001	.359
	Control	3.85 (0.51)	3.71 (0.49)	-0.14			
Boredom (ABSP)	Experimental	4.11 (0.61)	1.84 (0.52)	-2.27	92.11	< .001	.449
	Control	4.05 (0.58)	3.96 (0.55)	-0.09			

Thematic analysis of the semi-structured interviews clarified how the structural components of the platform satisfied the core tenants of Self-Determination Theory (autonomy, competence, and relatedness), effectively counteracting classroom boredom and speech apprehension.

"In traditional pronunciation drillings, I felt like a robot just repeating words, which was incredibly boring. But the digital quests allowed my guild to choose our own paths and scenarios. That freedom of choice made us feel in control of our learning, and the points system made us want to practice the difficult sounds over and over without getting tired." (Participant EG-19)

"The leaderboards and badges did not make me stressed; instead, they made the class feel like a cooperative game. Because we were working in 'Guilds' to solve intercultural communicative challenges, the pressure of being individually judged by the teacher disappeared. Our relatedness increased because we supported each other's pronunciation development to win the team quest." (Participant EG-52)

"I used to experience extreme anxiety whenever I had to speak English out loud. The gamified platform reframed errors as temporary setbacks rather than permanent academic failures. Getting a badge like 'Stress Navigator' after failing a task three times gave me a strong sense of competence and encouraged me to keep adjusting my pitch and intonation." (Participant EG-03)

Discussions

The empirical findings of this investigation provide robust verification for the integration of theory-driven gamification within tertiary EFL pronunciation instruction. The substantial increase in speech intelligibility and comprehensibility observed in the experimental group directly aligns with the assertions of Munro and Derwing (1995) and Levis (2011)^[14, 22], who argued that instructional

energy should be directed toward optimizing functional communicative clarity rather than chasing an elusive, native-like ideal. The digital quests and automated feedback systems offered an optimized instructional environment where learners could repeatedly test and refine core phonological features, such as final consonant articulation and lexical stress placement. This targeted practice successfully corrected the pronunciation-spelling ambiguities that frequently trouble Vietnamese university students.

Crucially, the study demonstrates that linguistic growth does not have to occur in a socio-cultural vacuum. By incorporating intercultural role-play simulations into a gamified framework, this study successfully bridged the gap between mechanical phonology and pragmatic reality. The significant positive shifts across all dimensions of the ICCAAS indicate that exposing learners to diverse phonetic variations within game-based problem-solving contexts actively builds intercultural empathy and linguistic flexibility. Rather than viewing regional accents through a rigid deficit model, learners developed an understanding of the English as a Lingua Franca paradigm, recognizing that language variations are natural characteristics of globalized communication. This finding advances the work of Jenkins (2007)^[10] and Yazan (2015) by demonstrating a concrete pedagogical pathway through which socio-cultural awareness can be seamlessly integrated into regular university coursework.

From a psychological perspective, the sharp reductions in classroom anxiety and academic boredom can be directly attributed to how well the gamified framework supported the core needs of Self-Determination Theory. Traditional pronunciation methods often alienate learners by relying on repetitive, decontextualized drilling that lowers self-efficacy and induces severe communication apprehension. In contrast, the quest-based architecture satisfied the students'

need for autonomy by offering meaningful learning choices, while the team-based "Guilds" fostered a sense of relatedness that transformed the classroom into a supportive community. This supportive environment effectively lowered the affective filter, turning potential performance anxiety into productive, goal-oriented engagement. Reframing pronunciation errors as iterative data points within a game loop rather than permanent academic marks allowed students to build the psychological resilience needed for long-term competence.

The pedagogical implications of these findings are highly significant for EFL curriculum design, educational technology development, and language policy in resource-constrained contexts. First, institutional language programs should systematically move away from native-centric evaluative standards, officially adopting an ELF orientation that prioritizes mutual intelligibility and intercultural pragmatics. Second, curriculum developers should reject traditional presentation-practice models in favor of blended learning formats that embed collaborative game mechanics, simulated cross-cultural interactions, and immediate, non-threatening feedback loops directly into the instructional design. Finally, for institutions operating in environments where access to native-speaking instructors is limited, these findings show that theory-driven, mobile-assisted gamification can serve as a highly effective, scalable equalizer. This approach cultivates autonomous, culturally empathetic, and communicatively capable language learners who are thoroughly equipped to navigate the linguistic diversity of modern international networks.

Conclusion

This study investigated the impact of an 8-week, Self-Determination Theory-grounded, gamified pronunciation instructional framework on Vietnamese university EFL learners' speech intelligibility, comprehensibility, and Intercultural Communicative Competence. The quantitative and qualitative findings conclusively show that integrating purposeful game mechanics with authentic intercultural simulations significantly outperforms traditional teacher-centered drilling methodologies. The gamified framework not only accelerated the acquisition of clear segmentals and suprasegmentals, but also dismantled native-centric biases, encouraging a broader perspective of global linguistic empathy toward World Englishes. Furthermore, by structurally satisfying the core psychological needs of autonomy, competence, and relatedness, the intervention effectively neutralized classroom boredom and speech anxiety, replacing them with sustained learner engagement. Ultimately, this research confirms that modern pronunciation pedagogy can successfully balance linguistic accuracy with abstract socio-cultural growth, offering a validated design blueprint for preparing modern learners for successful global communication.

Despite its positive outcomes, several limitations should be considered when interpreting the generalizability of this study. First, the investigation relied on a convenience sampling technique drawing from intact classes at a single public university in southern Vietnam, which may not fully represent the diverse student demographics across different geographic regions or institutional profiles. Second, the 8-week intervention period, while intensive, was relatively brief; this constrains the ability to draw conclusions regarding the long-term retention of phonological gains or

the permanent stabilization of intercultural attitudes over time. Third, the speech intelligibility measures were evaluated based on laboratory-recorded paragraph readings and structured picture descriptions, which may not completely capture the fluid, dynamic adjustments that occur during unscripted, real-world intercultural speech production. Finally, the automated scoring engine utilized within the mobile interface occasionally exhibited technical sensitivities when interpreting highly idiosyncratic L2 speech variations, introducing minor inconsistencies in initial individual feedback loops.

Based on these findings and limitations, several directions are proposed for future research in computer-assisted pronunciation training and intercultural pragmatics. First, future studies should implement longitudinal designs featuring delayed post-tests administered six months to one year post-intervention to rigorously evaluate the long-term retention of phonological intelligibility and the stability of ideological shifts regarding World Englishes. Second, researchers should expand the participant pool using randomized stratified sampling methods across multiple geographical contexts and different baseline proficiency levels to enhance the external validity of the gamified framework. Third, future pedagogical designs should incorporate advanced, real-time conversational AI models capable of generating adaptive, unscripted intercultural interactions to measure comprehensibility under natural communicative pressure. Finally, future work should conduct comparative investigations to isolate the specific motivational impacts of individual game elements such as comparing cooperative guild tasks against purely competitive leaderboards to determine the optimal configuration for reducing communication anxiety across diverse cultural learning styles.

References

1. Agustina T, Hajimia H, Yusran G, Halawiyah I, Farizi MI, Ariawan S. EFL Students' Perceptions and Experiences on the Utilization of Ling App in Learning Pronunciation. *Jurnal Pendidikan Progresif*, 2025;15(1):659-670. <https://doi.org/10.23960/jpp.v15i1.pp659-670>
2. Barcomb M, Cardoso W. Rock or Lock? Gamifying an online course management system for pronunciation instruction. *Calico Journal*, 2020;37(2):127-147. <https://utppublishing.com/doi/abs/10.1558/cj.36996>
3. Bøhn H, Hansen T. Assessing pronunciation in an EFL context: Teachers' orientations towards nativeness and intelligibility. *Language Assessment Quarterly*, 2017;14(1):54-68. <https://doi.org/10.1080/15434303.2016.1256407>
4. Bouchhioua N. The effects of explicit pronunciation instruction on the comprehensibility and intelligibility of Tunisian EFL learners. *International Journal of Research Studies in Language Learning*, 2017;6(3):73-88.
5. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 2006;3(2):77-101. <https://doi.org/10.1191/1478088706qp063oa>
6. Chen HC. In-Service Teachers' Intelligibility and Pronunciation Adjustment Strategies in English Language Classrooms. *English Language*

- Teaching,2016:9(4):30-53. <http://dx.doi.org/10.5539/elt.v9n4p30>
7. Deci EL, Ryan RM. The what and why of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*,2000:11(4):227-268. https://doi.org/10.1207/S15327965PLI1104_01
 8. Galante A, Piccardo E. Teaching pronunciation: toward intelligibility and comprehensibility. *ELT journal*,2022:76(3):375-386.
 9. Horwitz EK, Horwitz MB, Cope J. Foreign language classroom anxiety. *The Modern language journal*,1986:70(2):125-132.
 10. Jenkins J. *English as a lingua franca: Attitude and identity*. Oxford University Press, 2007.
 11. Jeong H, Lindemann S. Beyond ideologies of nativeness in the intelligibility principle for L2 English pronunciation: A corpus-supported review. *System*, 2025, 129, 103599. <https://doi.org/10.1016/j.system.2025.103599>
 12. Khelef H, Khelef A. Gamification For Pronunciation: How To Energize Learning?. *Ilkogretim Online*,2024:23(4):888-901.
 13. Levis JM. Changing contexts and shifting paradigms in pronunciation teaching. In *Pronunciation*. Routledge, 2005, 265-272.
 14. Levis J, LeVelle K. Pronunciation and intelligibility: An overview of the conference. *Pronunciation in Second Language Learning and Teaching Proceedings*, 2011, 2(1). <https://www.iastatedigitalpress.com/psllt/article/id/15155/>
 15. Luu DP, Le TTV. Exploring the relationship between vocabulary learning problems and strategies among non-English major students at a public university. *HO CHI MINH CITY OPEN UNIVERSITY JOURNAL OF SCIENCE-SOCIAL SCIENCES*, 2026.
 16. Luu DP, Nguyen HMA, Nguyen GH. Internal and external determinants of English majors' attitudes towards pronunciation learning at a Vietnamese public university. *Thu Dau Mot University Journal of Science*,2025:7(4):813-833. <https://doi.org/10.37550/tdmu.EJS/2025.04.679>
 17. Luu DP, Nguyen VTX, Do LC, Nguyen TP. A Self-Determination Theory Model of Gamified EFL Intrinsic Motivation. *EuroGlobal Journal of Linguistics and Language Education*,2025:2(5):60-74. <https://doi.org/10.69760/egjle.2505005>
 18. Luu PD. EFL students' Intercultural Willingness to Communicate through Intergenerational Storytelling. *Journal of Philology and Educational Sciences*,2026:5(1):23-44. <https://doi.org/10.53898/jpes2026512>
 19. Luu PD, Nguyen PT. Exploring Common Pronunciation Errors Faced by English Majors at a Vietnamese Public University. *Journal of Foreign Language Teaching and Learning*,2026:11(1):194-228. <https://doi.org/10.18196/ftl.v11i1.29346>
 20. Luu PD, Anh NHM, Truc NDT, Dung LN, Vi LTT, Thu DTX. The Socio-Pragmatics of Digital Slang in Post-Pandemic Online Communities. *Acta Globalis Humanitatis et Linguarum*,2025:2(5):28-39. <https://doi.org/10.69760/aghel.0250050003>
 21. Luu PD, Nguyen HMA, Nguyen DTT, Le ND, Le TTV, Do TXT. Integrating gamification into the classroom: Theoretical mechanisms for enhancing student engagement. *International Journal of Academic Research and Writing*,2025:7(4):23-30.
 22. Munro MJ, Derwing TM. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*,1995:45(1):73-97.
 23. Pekrun R. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*,2006:18(4):315-341. <https://doi.org/10.1007/s10648-006-9029-9>
 24. Saito K. Experienced teachers' perspectives on priorities for improved intelligible pronunciation: The case of Japanese learners of English. *International Journal of Applied Linguistics*,2014:24(2):250-277. <https://doi.org/10.1111/ijal.12026>
 25. Sholekhah MF, Fakhurriana R. The use of ELSA speak as a Mobile-Assisted Language Learning (MALL) towards EFL students' pronunciation. *JELITA: Journal of Education, Language Innovation, and Applied Linguistics*,2023:2(2):93-100.
 26. Shrestha A, Rajbhandari Y, Khadka N, Bista A, Marahatta A, Dahal R, et al. Status of micro/mini-grid systems in a himalayan nation: a comprehensive review. *IEEE Access*,2020:8:120983-120998.
 27. Shroff R, Keyes C. Gamified Pedagogy: Examining how a Phonetics App Coupled with Effective Pedagogy can Support Learning. *Electronic Journal of Foreign Language Teaching*, 2020, 17(1). <https://e-flt.nus.edu.sg/wp-content/uploads/2020/09/shroff.pdf>