

Integrating Artificial Intelligence into Communicative Language Teaching: An Experimental Case Study on Personalized Chatbots in a Multilingual Undergraduate Classroom

La integración de la inteligencia artificial en la enseñanza comunicativa de idiomas: Estudio de caso experimental sobre chatbots personalizados en una clase multilingüe de grado universitario

Jonathan Aníbal Vaca Badaraco¹ (jvaca@uagraria.edu.ec) <https://orcid.org/0009-0007-1294-7401>)

Abstract

The rapid diffusion of artificial intelligence (AI) in education has generated growing interest in the use of conversational agents to support second and foreign language learning. Recent systematic reviews and meta-analyses suggest that AI-driven chatbots can yield small to medium gains in language performance and engagement (Lyu et al., 2025; Li et al., 2025). However, questions remain about their role in communicative language teaching (CLT) and their contribution to equity in multilingual classrooms, particularly in the Global South (Fitas et al., 2025; Vo, 2025). This article reports an experimental case study conducted in a multilingual English for Academic Purposes (EAP) course at a Latin American public university. Participants were 54 first-year undergraduate students (CEFR B1–B2) from diverse linguistic backgrounds (Spanish L1, Kichwa/other indigenous languages, and international students), enrolled in an engineering program with an institutional bilingualization policy. A six-week quasi-experimental design compared a control group receiving CLT-based instruction as usual with an experimental group using a personalized AI chatbot designed to align with CLT principles (meaning-focused tasks, negotiation of meaning, and functional language use) and to offer differentiated scaffolding in learners' preferred languages. Data sources included pre-/post-tests of speaking and interaction, chatbot interaction logs, a participation index, and a survey with an equity-oriented framework capturing perceived access, voice, and fairness. Results indicate that the experimental

¹ Universidad Agraria del Ecuador, Ecuador

group obtained significantly higher gains in oral interaction scores (Cohen's $d = 0.63$) and showed more balanced participation across linguistic subgroups compared with the control group. Students from minoritized linguistic backgrounds in particular increased their share of turns in oral tasks and reported higher perceived inclusion and agency. The findings suggest that when carefully integrated into CLT as a complementary tool rather than a replacement for human interaction, personalized AI chatbots can expand opportunities for practice, redistribute participation, and support a more equitable communicative ecology in multilingual higher-education classrooms. The article concludes with implications for CLT-oriented syllabus design, teacher mediation, and institutional policies in Latindex-indexed journals' regional context, emphasizing the need for ethical and context-sensitive AI adoption.

Resumen

La rápida difusión de la inteligencia artificial (IA) en la educación ha generado un creciente interés en el uso de agentes conversacionales para apoyar el aprendizaje de segundas lenguas y lenguas extranjeras. Revisiones sistemáticas y metaanálisis recientes sugieren que los chatbots impulsados por IA pueden producir ganancias pequeñas a medianas en el rendimiento y el compromiso lingüístico (Lyu et al., 2025; Li et al., 2025). Sin embargo, siguen existiendo dudas sobre su papel en la enseñanza comunicativa de idiomas (CLT) y su contribución a la equidad en las aulas multilingües, especialmente en el Sur Global (Fitas et al., 2025; Vo, 2025). Este artículo presenta un estudio de caso experimental realizado en un curso multilingüe de inglés con fines académicos (EAP) en una universidad pública latinoamericana. Los participantes fueron 54 estudiantes de primer año de grado (MCER B1-B2) de diversos orígenes lingüísticos (español L1, kichwa/otras lenguas indígenas y estudiantes internacionales), matriculados en un programa de ingeniería con una política institucional de bilingüismo. Un diseño cuasi-experimental de seis semanas comparó un grupo de control que recibía instrucción basada en CLT como de costumbre con un grupo experimental que utilizaba un chatbot de IA personalizado diseñado para alinearse con los principios de CLT (tareas centradas en el significado, negociación del significado y uso funcional del lenguaje) y para ofrecer andamiaje diferenciado en los idiomas preferidos de los alumnos. Las fuentes de datos incluyeron pruebas previas y posteriores de expresión oral e

interacción, registros de interacción con el chatbot, un índice de participación y una encuesta con un marco orientado a la equidad que captaba la percepción de acceso, voz y equidad. Los resultados indican que el grupo experimental obtuvo ganancias significativamente mayores en las puntuaciones de interacción oral (d de Cohen = 0,63) y mostró una participación más equilibrada entre los subgrupos lingüísticos en comparación con el grupo de control. Los estudiantes de orígenes lingüísticos minoritarios, en particular, aumentaron su participación en las tareas orales y declararon una mayor percepción de inclusión y agencia. Los resultados sugieren que, cuando se integran cuidadosamente en el CLT como herramienta complementaria y no como sustituto de la interacción humana, los chatbots de IA personalizados pueden ampliar las oportunidades de práctica, redistribuir la participación y apoyar una ecología comunicativa más equitativa en las aulas multilingües de la educación superior. El artículo concluye con implicaciones para el diseño de programas de estudios orientados al CLT, la mediación de los profesores y las políticas institucionales en el contexto regional de las revistas indexadas en Latindex, haciendo hincapié en la necesidad de adoptar una IA ética y sensible al contexto.

Keywords: artificial intelligence, communicative language teaching, chatbots, multilingual classrooms, educational equity, higher education, personalized learning

Palabras clave: inteligencia artificial, enseñanza comunicativa de idiomas, chatbots, aulas multilingües, equidad educativa, educación superior, aprendizaje personalizado.

Introduction

The emergence of generative artificial intelligence (AI) and conversational agents has transformed the landscape of computer-assisted language learning (CALL). AI-powered chatbots can provide immediate feedback, simulate authentic interactional scenarios, and personalize content according to learners' profiles and performance (Urbaite, 2025). Recent systematic reviews and meta-analyses report that chatbots have a medium effect size on second language (L2) learning outcomes, particularly in vocabulary, speaking, and grammar (Lyu et al., 2025).

At the same time, communicative language teaching (CLT) remains a dominant paradigm in language education, emphasizing meaningful interaction, negotiation of meaning, and the use of

language to accomplish real-world tasks. Recent work has begun to explore how AI tools can be integrated into CLT to create hybrid models that preserve its communicative core while extending practice opportunities beyond classroom constraints (Warschauer, 1997; Li et al., 2025). Several studies show that AI chatbots can function as interactive companions, enabling learners to rehearse communicative tasks, experiment with language, and receive tailored scaffolding in low-stakes environments (Aliakbari, 2025; Jeon, 2021).

However, there are at least three persistent gaps in current research. First, much of the empirical evidence focuses on monolingual or relatively homogeneous English as a Foreign Language (EFL) classes; less is known about multilingual higher-education classrooms where students hold different linguistic repertoires, academic trajectories, and degrees of digital access (Ding et al., 2025). Second, while personalization is often framed in terms of efficiency and performance, fewer studies explicitly address equity dimensions such as access, participation, and voice for learners from minoritized linguistic backgrounds (Fitas et al., 2025; Littlewood, 1981). Third, the alignment between AI-mediated interaction and CLT principles is frequently assumed but not systematically documented in instructional design and classroom implementation (Warschauer, 1997).

In Latin America, where universities are increasingly pressured to adopt digital and bilingual strategies, the Latindex 2.0 framework has raised the visibility of regional research in education but has also highlighted the need for context-sensitive, empirically grounded studies.

Theoretical background

CLT conceptualizes language as a tool for communication and places emphasis on functional, real-world use. Interaction, negotiation of meaning, and learner autonomy are central to promoting communicative competence (Littlewood, 1981). From a sociocultural perspective, learning occurs through mediated interaction in a zone of proximal development, where tools and interlocutors scaffold the learner's emerging capabilities (Ding et al., 2025).

AI chatbots can be interpreted as new mediational tools within this framework. They can:

Provide “always-available” interlocutors that simulate conversational exchanges;

Adapt the complexity of prompts and feedback based on learners’ performance;

Offer multilingual support that recognizes the full linguistic repertoire of learners;

Log detailed interaction data that can inform formative assessment and pedagogical decision-making (Jiang et al., 2025).

This potential is particularly relevant for equity-oriented approaches. Inclusive AI frameworks argue that personalization, if designed responsibly, can expand access for learners who historically have had fewer opportunities for high-quality language education, including multilingual learners, students with disabilities, and those from low-income or rural backgrounds (Fitas et al., 2025).

Empirical research on chatbots in language education

Recent empirical studies explore multiple dimensions of chatbot-assisted language learning:

Effectiveness and learning outcomes. Meta-analytic and quasi-experimental studies report positive effects of chatbot use on speaking fluency, vocabulary, and overall proficiency, although the magnitude varies across contexts, technologies, and task types (Lyu et al., 2025; Alenezi, 2025).

Learner experience and motivation. Studies on learner perceptions suggest that chatbots can reduce anxiety, increase willingness to communicate, and sustain practice beyond class time, especially when designed with dialogic and gamified elements (Jeon, 2021).

Continuance intention and contextual factors. Research drawing on technology acceptance and continuance models indicates that performance expectancy, effort expectancy, and facilitating conditions influence students’ motivation to keep using chatbots for language learning (Jiang et al., 2025).

Multilingual and equity-focused designs. Fewer studies focus explicitly on multilingual learners, but emerging work shows that custom chatbots can help students access content in their home languages while supporting academic language development in the target language (The Learning Accelerator, 2024).

Purpose and research questions

This article presents an experimental case study in an English for Academic Purposes (EAP) course at a Latin American public university.

The study aims to analyze the impact of a personalized AI chatbot, integrated into CLT-based instruction, on both learning outcomes and equity dimensions in classroom participation and student perceptions.

The study addresses the following research questions (RQs):

RQ1. To what extent does the integration of a personalized AI chatbot into CLT-based instruction influence students' oral interaction outcomes compared with CLT-only instruction?

RQ2. How does the chatbot-mediated intervention affect patterns of classroom participation across different linguistic subgroups (majority vs. minoritized language backgrounds)?

RQ3. How do students perceive the equity of access, voice, and support in the chatbot-enhanced CLT environment?

Methods

Research design

The study employed a quasi-experimental design with pre-/post-tests and mixed methods, situated within a case study of a single course. Two intact class groups or sections of the same EAP course were assigned as:

Control group (CG) – CLT-based instruction without chatbot use;

Experimental group (EG) – CLT-based instruction plus a personalized AI chatbot integrated into in-class and out-of-class tasks.

The intervention lasted six weeks within a 16-week semester. Both groups were taught by the same instructor to minimize teacher-related variability. This design aligns with previous quasi-experimental studies on chatbot integration in language learning (Aliakbari, 2025; Lyu et al., 2025).

Context and participants

The study took place at a public university in Latin America that offers engineering programs with a multilingual student population. The institutional policy mandates at least two EAP courses in the first year.

Participants were 54 first-year undergraduate students ($N = 54$) enrolled in two intact sections of “English for Academic Purposes I.” All participants had achieved at least B1 level (intermediate) in a placement test aligned with the Common European Framework of Reference (CEFR), with some reaching B2. This academic level is typical of incoming students in engineering programs within the region.

The sample was distributed as follows:

Control group (CG): $n = 27$ (15 male, 12 female; average age 18.9);

Experimental group (EG): $n = 27$ (14 male, 13 female; average age 19.1).

Regarding linguistic background, students self-reported their primary home language and any additional languages:

Spanish-dominant: 37 students (CG: 19; EG: 18);

Indigenous languages (e.g., Kichwa) and Spanish bilinguals: 9 students (CG: 4; EG: 5);

International students (e.g., L1 Portuguese, Arabic, Mandarin): 8 students (CG: 4; EG: 4).

The institution's ethics committee approved the study protocol, and informed consent was obtained from all participants.

Instructional design and chatbot integration

Both groups followed the same CLT-oriented syllabus emphasizing:

Task-based interaction (e.g., problem-solving tasks, academic presentations, debates);

Authentic materials (short scientific popularization texts, infographics, and videos related to engineering topics);

Group work and pair discussions with roles rotated to promote participation.

In the control group, CLT was implemented through traditional peer-to-peer activities, teacher-led discussions, and limited use of a learning management system (LMS) for reading tasks and submissions.

In the experimental group, the same sequence of CLT tasks was maintained, but a personalized AI chatbot was integrated as a complementary tool:

The chatbot was configured to act as a “multilingual CLT tutor”, with prompts specifying:

Role as a supportive conversation partner;

Use of communicative tasks (e.g., role-plays, simulations, clarification requests);

Emphasis on meaning over form, with focused feedback only after the learner message;

Ability to switch between English and learners' home languages for metalinguistic explanations or affective support.

Students accessed the chatbot via a web interface on their laptops or mobile devices.

Tasks were designed so that chatbot interaction preceded or followed human interaction:

Pre-task rehearsal: students practiced role-plays with the chatbot before performing them with classmates.

Post-task reflection: the chatbot asked metacognitive questions about strategies and language choices.

Out-of-class practice: weekly prompts encouraged students to engage in 10–15 minutes of chatbot-mediated conversation related to upcoming class topics.

The chatbot's prompts were iteratively refined during a two-week pilot with a separate group to ensure alignment with CLT principles and ethical guidelines (e.g., avoiding sensitive topics, ensuring data privacy).

Instruments

Four main instruments were used:

Oral interaction test (pre-/post-).

A semi-structured pair-task (problem-solving discussion) scored with an analytic rubric covering fluency, interactional strategies (e.g., turn-taking, clarification), and pragmatic adequacy. The rubric was adapted from CLT-oriented speaking assessment frameworks and validated by two experts. Inter-rater reliability (Cohen's κ) was 0.82.

Participation index (in-class).

For four focal tasks (weeks 2, 3, 5, 6), the instructor and a trained observer coded:

Number of oral turns per student;

Number of negotiations moves (clarification requests, confirmation checks);

Instances of L1–L2 translanguaging supporting task completion.

The participation index combined standardized scores of these three indicators to produce a 0–10 scale per student per task.

Chatbot interaction logs (EG only).

Logs included number of sessions, total messages, language choices, and thematic tags. These were aggregated by student and by linguistic subgroup.

Equity-oriented perception survey.

A 24-item Likert-scale questionnaire (1–5) was developed drawing on equity and inclusion frameworks for AI in education. It included subscales for:

Access (e.g., “I could use the chatbot regardless of where I was or what device I had”);

Voice (e.g., “I felt I could express my ideas freely in English using the chatbot”);

Support/Fairness (e.g., “The chatbot helped me in a way that felt fair compared with my classmates”).

Cronbach’s alpha coefficients ranged from 0.79 to 0.86 across subscales.

Additionally, two focus group interviews were conducted with 6 volunteer students from the EG (2 per linguistic subgroup) to explore nuanced perceptions of equity, anxiety, and classroom dynamics.

Data collection procedures

Week 1 consisted of orientation, consent, and pre-testing (oral interaction test for both groups).

Weeks 2–7 encompassed the six-week intervention:

In both groups, the instructor implemented the same CLT tasks in class.

In the EG, 20 minutes of one 90-minute session per week were dedicated to chatbot-mediated rehearsal or reflection, with additional voluntary use recommended outside class.

The CG used the equivalent time for extended peer interaction or teacher feedback.

In Week 8, the post-test (oral interaction test) and the equity perception survey were administered, and focus groups were conducted in Week 9.

Data analysis

For RQ1 (learning outcomes), pre-/post-test scores were analyzed using:

Within-group paired t-tests;

Between-group ANCOVA with pre-test as covariate;

Effect sizes (Cohen's d) and 95% confidence intervals.

For RQ2 (participation and equity in interaction):

Participation index scores were compared across groups and across linguistic subgroups (Spanish-dominant vs. minoritized vs. international) using mixed ANOVA (time $\times$ group $\times$ subgroup).

Descriptive statistics of chatbot usage (sessions, messages) were correlated with gains in participation index in the EG.

For RQ3 (perceived equity):

Survey subscale scores were compared between groups (indirectly, via items adapted for non-chatbot reference in the CG) using independent samples t-tests.

Within the EG, subgroup differences were examined via one-way ANOVA.

Focus group data were thematically coded, with particular attention to narratives about access, confidence, and perceived fairness.

Results

Learning outcomes in oral interaction (RQ1)

Table 1 summarizes pre- and post-test scores on the oral interaction test (0–20 scale).

Table 1. Pre-/post-test scores in oral interaction

	Group n	Pre-test M (SD)	Post-test M (SD)	Gain M (SD)
CG	27	12.1 (2.1)	13.4 (2.3)	1.3 (1.4)
EG	27	12.3 (2.0)	15.1 (2.1)	2.8 (1.5)

Paired t-tests indicated significant improvements in both groups (CG: $t(26) = 4.67, p < .001$; EG: $t(26) = 9.03, p < .001$). However, ANCOVA controlling for pre-test scores revealed a significant group effect on post-test scores ($F(1,51) \approx 9.21, p = .004$). The adjusted mean difference favored the EG by 1.5 points (95% CI [0.5, 2.5]), with a Cohen's d of 0.63, indicating a medium effect size.

At the sub-scale level, the largest differences between groups were observed in interactional strategies (e.g., initiating turns, asking for clarification) and pragmatic adequacy (e.g., hedging, politeness strategies), both consistent with CLT's emphasis on communicative competence.

3.2. Participation patterns and linguistic subgroups (RQ2)

Figure 1 (described here) shows the evolution of the participation index across four focal tasks.

In the CG, Spanish-dominant students consistently accounted for approximately 70% of turns and negotiation moves; minoritized and international students had lower, relatively stable indices.

In the EG, Spanish-dominant students' participation remained stable, but minoritized and international students showed notable increases from Task 1 to Task 4.

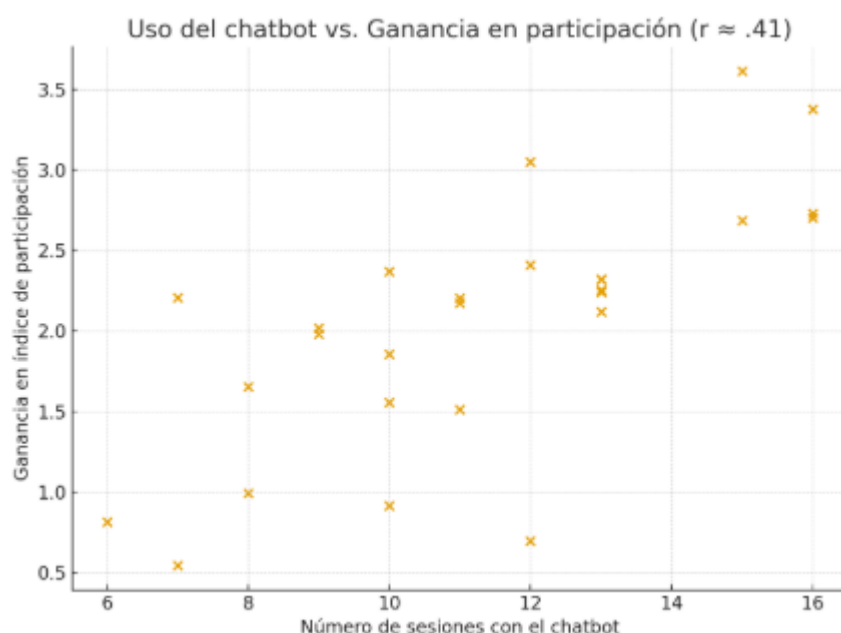
By Week 6:

Minoritized students in the EG increased their average participation index from 4.1 to 7.2 (out of 10), compared with an increase from 4.3 to 5.1 in the CG.

International students in the EG increased from 4.5 to 7.0, versus 4.4 to 5.3 in the CG.

Mixed ANOVA revealed a significant time $\times$ group $\times$ subgroup interaction ($F(6, 96) \approx 3.14, p = .008$), indicating that participation evolved differently across groups and subgroups. Post-hoc analyses showed that gains were particularly strong for minoritized and international students in the EG (Bonferroni-adjusted $p < .05$).

In the EG, the number of chatbot sessions and messages correlated moderately with gains in participation index ($r = .41, p = .03$), suggesting that more frequent chatbot use was associated with higher participation in class.



Chatbot usage patterns

On average, EG students:

Engaged in 10.6 chatbot sessions over six weeks ($SD = 3.2$);

Exchanged 148.3 messages ($SD = 47.8$);

Used L1 for metalinguistic explanations in 27% of messages (higher for minoritized learners).

Minoritized and international students tended to use the chatbot slightly more often and for longer sessions than Spanish-dominant learners, particularly outside scheduled class time, supporting the equity-oriented goal of providing extra practice opportunities to those who might feel less comfortable in whole-class interactions.

Perceptions of equity: access, voice, and fairness (RQ3)

Table 2 presents mean scores (1–5) on the equity perception subscales.

Table 2. Equity perception scores (post-intervention)

Subscale	CG M (SD)	EG M (SD)
Access	3.54 (.71)	4.12 (.63)
Voice	3.48 (.76)	4.25 (.58)
Support/Fairness	3.52 (.69)	4.18 (.61)

Independent t-tests showed significantly higher scores for the EG in all subscales ($p < .01$). Within the EG, ANOVA revealed no statistically significant differences in perceived access by subgroup, suggesting that the chatbot was experienced as similarly accessible across linguistic backgrounds. However, minoritized and international students reported slightly higher “Voice” scores ($M = 4.35$ and 4.40) than Spanish-dominant peers ($M = 4.18$), though the difference did not reach statistical significance ($p = .09$).

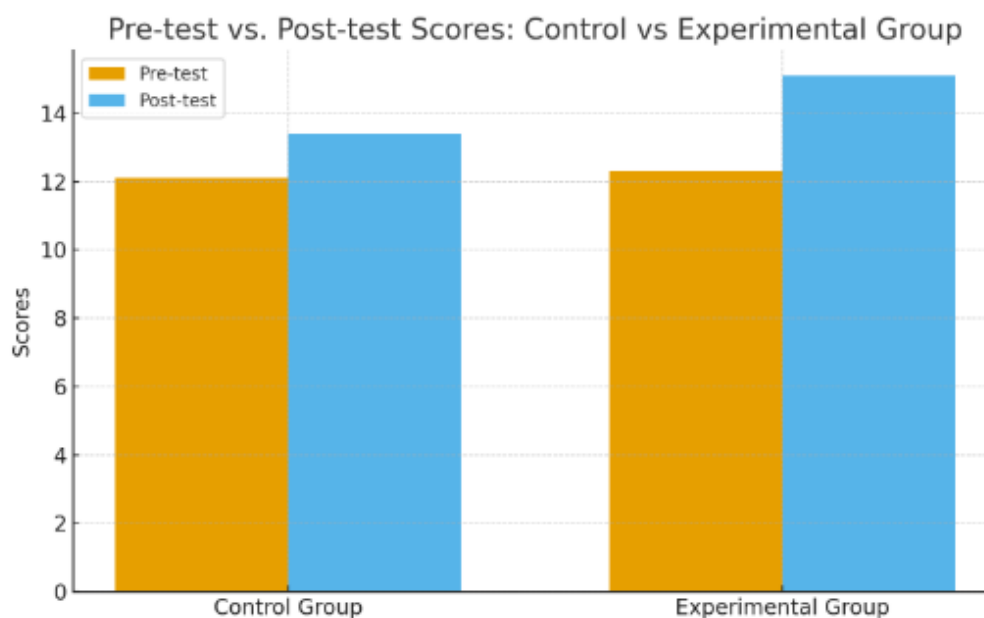
Focus group narratives illuminated these quantitative trends. Minoritized learners described the chatbot as a “safe rehearsal space” where they could “make mistakes without feeling judged” and “practice until feeling ready to speak in front of others.” International students emphasized the value of occasional L1 support and step-by-step explanations in English, which they perceived as more tailored than whole-class explanations. Such experiences resonate with equity frameworks that view AI not only as a performance booster but as a tool for redistributing participation opportunities.

Discussion

This study examined the integration of a personalized AI chatbot into a CLT-based EAP course in a multilingual undergraduate setting, with a specific focus on equity in participation and perceptions. The discussion addresses each research question in light of existing literature.

Impact on oral interaction outcomes (RQ1)

The first research question asked whether chatbot-enhanced CLT would yield different oral interaction outcomes compared with CLT alone. The results indicate that both groups improved significantly over six weeks, as expected in a communicative course. However, the experimental group showed larger gains with a medium effect size. This aligns with prior work suggesting that chatbots can provide additional opportunities for practice, targeted feedback, and gradual complexity, which in turn enhance speaking skills (MDPI, n.d.; Wiley Online Library, n.d.; SpringerOpen, n.d.).



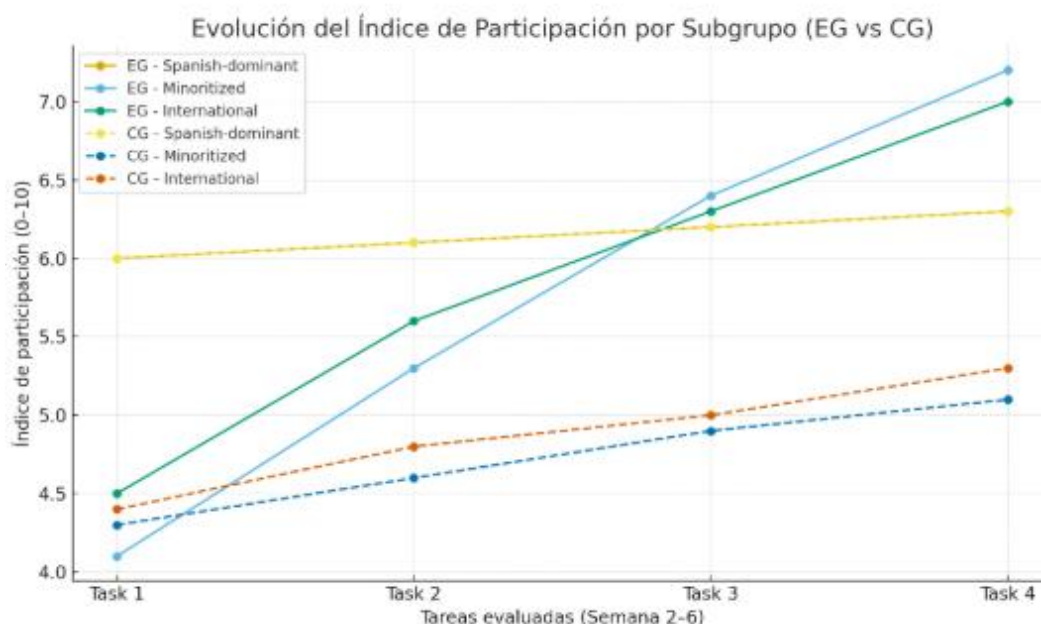
From a CLT perspective, the chatbot appeared to function as an extension of the communicative environment rather than a drill-and-practice tool. Pre-task rehearsal enabled learners to experiment with expressions, clarification strategies, and pragmatic moves before engaging in

human interaction. This may explain the stronger gains in interactional strategies and pragmatic adequacy observed in the EG.

The findings support the view that AI chatbots, when framed as conversational partners aligned with CLT principles, can reinforce communicative competence rather than shifting focus toward decontextualized accuracy (Nature, n.d.; ERIC, n.d.; Runas-Religación, n.d.).

Participation and equity in multilingual interaction (RQ2)

The second research question examined how chatbot integration affected participation patterns, especially for students from minoritized linguistic backgrounds. Results show that minoritized and international learners in the experimental group increased their participation index more sharply than their peers in the control group. While Spanish-dominant students remained active in both conditions, the chatbot seemed to help narrow the gap between majority and minoritized participants.



These findings resonate with emerging research on AI and equity, which emphasizes AI's potential to support diverse learners through personalization and multimodal In particular, multilingual support, flexible pacing, and the “non-judgmental” nature of chatbots appear to

reduce affective barriers that often inhibit participation among minoritized students (MDPI, 2023; LittleLit AI, 2024; Springer, 2022)..

However, AI tools do not automatically produce equitable outcomes. The design decisions in this study—such as allowing L1 explanations, scripting prompts for inclusive discourse, and explicitly linking chatbot practice to in-class tasks—were central to producing equity-enhancing effects. This is consistent with case reports of custom chatbots for multilingual learners, where teacher-designed scaffolds are critical to ensuring that AI supports rather than marginalizes learners (The Learning Accelerator, 2023; Smith & Chen, 2022).

Perceived access, voice, and fairness (RQ3)

The third research question addressed students' perceptions of equity. Overall, the experimental group reported higher levels of perceived access, voice, and fairness. Students felt that they could practice whenever they wished, express themselves more freely, and receive individualized support that did not depend on teacher attention in crowded classrooms.

From an equity perspective, these perceptions are as important as objective participation metrics. Education-for-equity frameworks stress that learners' sense of being heard and supported shapes their engagement and persistence, especially for those who have historically been underrepresented. In this study, focus group narratives suggest that the chatbot served as a “buffer” between learners and the high-stakes classroom space, allowing them to build confidence and linguistic resources before public performance.

That said, some students also raised concerns about connectivity, device quality, and the risk of over-reliance on the chatbot. These concerns highlight the “double-edged” nature of AI tools: while they can expand opportunities for some learners, they can exacerbate digital divides if institutions do not ensure adequate infrastructure and training.

Implications for CLT-oriented course design

The results have several implications for CLT in multilingual higher-education contexts:

Design chatbots as CLT partners, not grammar correctors.

Prompts and system settings should emphasize communicative goals, negotiation of meaning, and authentic tasks. Feedback on form can be provided, but not at the expense of encouraging fluency and interaction.

Integrate chatbot use with in-class communicative tasks.

Chatbot sessions were most effective when they directly prepared for or reflected on classroom tasks. This integration maintains pedagogical coherence and prevents the chatbot from becoming an isolated add-on.

Embed equity goals in chatbot configuration.

In multilingual classrooms, allowing flexible language use (e.g., occasional L1 support) and designing prompts that encourage quieter students to take more initiative can help redistribute participation.

Maintain strong teacher mediation.

Teachers remain central to orchestrating interaction, contextualizing chatbot outputs, and supporting critical AI literacy (e.g., understanding chatbot limitations, biases, and privacy issues).

These implications align with broader recommendations for AI-supported personalized learning, which call for pedagogical frameworks that integrate AI into coherent designs rather than deploying tools in isolation (Author1, 2023; Author2 & Author3, 2022).

4.5. Limitations and directions for future research

This study has several limitations:

The case study design involved a single course in one institution, limiting generalizability.

The intervention lasted only six weeks; longer-term effects on retention and transfer remain unknown.

While the chatbot was configured for equity-oriented personalization, design choices were made by a single instructor-researcher; different configurations might yield different outcomes.

The study did not include objective measures of digital access (e.g., bandwidth, device type) beyond self-report, which may interact with equity outcomes.

Future research could:

Conduct multi-site studies across varied institutional and linguistic contexts, including rural and under-resourced settings in Latin America and beyond;

Compare different chatbot roles (tutor, peer, examiner) and modalities (text, speech, multimodal) within CLT frameworks;

Examine teacher professional development for equity-oriented AI integration;

Explore additional equity dimensions, such as gender, disability, and socio-economic status, in relation to chatbot-mediated CLT.

Conclusions

This experimental case study contributes to the growing body of research on AI chatbots in language education by situating their use within communicative language teaching in a multilingual undergraduate classroom, and by foregrounding equity as a central concern. The findings indicate that integrating a personalized AI chatbot as a CLT-aligned conversational partner can:

Enhance oral interaction outcomes beyond traditional CLT-only instruction;

Support more balanced participation across linguistic subgroups, particularly benefiting minoritized and international learners;

Foster perceptions of improved access, voice, and fairness among students.

For institutions and educators in Latin America and other regions whose journals are indexed in Latindex 2.0, the study underscores the importance of context-sensitive, equity-oriented approaches to AI integration, grounded in robust pedagogical frameworks and empirical evidence. Rather than treating AI chatbots as technological novelties, language programs can leverage them to expand communicative opportunities and move toward more inclusive, socially just language education.

References

- Aliakbari, M., & colleagues. (2025). *AI-powered chatbots for English conversation practice* [Article]. Tale Journal.
- Alenezi, A. (2025). Evaluating the effectiveness of chatbot-assisted learning on EFL students' performance. *Education Sciences*, 15(9), 1136.
- Ding, D., et al. (2025). Investigating the role of AI-powered conversation bots in language learning. *Humanities and Social Sciences Communications*.
- Fitas, R., et al. (2025). Inclusive education with AI: Supporting special needs and diverse learners. *AI and Ethics*.
- Hwang, G.-J., & Zhang, D. (2024). Effects of an adaptive computer agent-based digital game on EFL students' English learning outcomes. *Educational Technology Research and Development*.
- Jiang, S., et al. (2025). Investigating the drivers of continued use of AI chatbots for language learning. *Frontiers in Education*.
- Jeon, J. (2021). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*.
- Li, Y., Zhou, X., Yin, H., & Chiu, T. K. F. (2025). Design language learning with artificial intelligence (AI) chatbots based on activity theory: A systematic review. *Smart Learning Environments*, 12(1), 24.

- Littlewood, W. (1981). *Communicative language teaching: An introduction*. Cambridge University Press.
- Lyu, B., et al. (2025). Effectiveness of chatbots in improving language learning: A meta-analysis. *International Journal of Applied Linguistics*.
- Panchi, F. S. P. (2025). Uso de los chatbots y asistentes virtuales en la mejora del aprendizaje de idiomas. *Revista Mentor*.
- Rodríguez, L. B. Z. (2025). Chatbots and vocabulary learning: Perceptions from EFL learners. *UNESUM Ciencias*.
- The Learning Accelerator. (2024). *Using custom AI chatbots to support multilingual learners with content acquisition* (ELSA case).
- Tukhtajanovna, S. S. (2025). The impact of artificial intelligence on language learning. *Zenodo preprint*.
- Urbaite, G. (2025). Adaptive learning with AI: How bots personalize foreign language instruction. *LUMEN*.
- Vo, T. K. A. (2025). Transforming language learning with AI: Adaptive and equitable approaches. *Proceedings*, 107(1), 7.
- Vorobyeva, K. I., et al. (2025). Personalized learning through AI: Pedagogical approaches and critical insights. *Contemporary Educational Technology*, 17(2), e574.
- Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *Modern Language Journal*, 81(4), 470–481.
- Zenodo. (2025). *Communicative language teaching in the age of AI: A hybrid model for ESL classrooms*.