UDK 81’243:004.85 378.014.3

Иновације у настави, XXXVIII, 2025/4, стр. 38–63

DOI: 10.5937/inovacije2504038T

Рад примљен: 4. 11. 2025.

Рад прихваћен: 30. 11. 2025.

Оригинални научни рад

Saida Tobbi[1]

Batna 2 University, Faculty of Letters and Foreign Languages, Batna, Algeria

AI-Mediated Oral Practice and the Dynamics of Anxiety, Performance, and Self-Regulation: A Mixed-Methods Comparison of Instructional Modes

Summary: *AI-mediated oral-practice tools offer scalable, low-stakes opportunities for second-language (L2) speaking practice, yet evidence of their pedagogical and affective impact remains limited. This quasi-experimental study investigated whether AI-mediated instruction enhances foreign-language anxiety (FLA), oral performance, and self-regulated learning (SRL) compared with blended and traditional human-taught instruction. A total of 150 learners (50 per group) from the English Department of Batna 2 University completed a six-week intervention integrating the AI pronunciation and fluency application ELSA Speak. Pre- and post-tests measured FLA, oral performance, and SRL. ANCOVAs controlling for baseline scores assessed adjusted group differences, and qualitative thematic analysis of reflective journals provided explanatory depth. Results revealed significant main effects of instructional mode on all outcomes: FLA ($F(2,146) = 32.47$, $p < .001$, partial $\eta^2 = .31$), oral performance ($F(2,146) = 23.68$, $p < .001$, partial $\eta^2 = .25$), and SRL ($F(2,146) = 31.02$, $p < .001$, partial $\eta^2 = .30$). AI-mediated learners reported markedly lower anxiety and higher performance and SRL than both comparison groups, with large standardized effect sizes (Hedges’ $g \approx 1.1–1.3$). Qualitative findings identified four interrelated themes—emotional regulation, self-regulatory growth, pronunciation and fluency development, and technological motivation—highlighting how supportive feedback and autonomous practice fostered engagement and confidence. These results demonstrate that AI-mediated oral practice can meaningfully reduce anxiety, strengthen self-regulation, and improve L2 speaking outcomes within a short intervention. The study advances understanding of AI as an active pedagogical agent and supports its thoughtful integration alongside human instruction.*

Keywords: *AI-mediated instruction; foreign-language anxiety, oral performance, self-regulated learning, computer-assisted language learning*

1 s.tobbi@univ-batna2.dz

https://orcid.org/0009-0005-9602-3411

Copyright © 2025 by the publisher Faculty of Education, University of Belgrade, SERBIA.

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original paper is accurately cited.

Introduction

The rapid advancement of AI is transforming English as a Foreign Language (EFL) pedagogy, particularly in the teaching of oral communication. Among the most influential innovations are AI-powered conversational agents—chatbots capable of sustaining interactive, context-sensitive dialogue with learners. Unlike early computer-assisted language learning (CALL) programs based on fixed drills, contemporary generative and voice-enabled tools such as *ChatGPT* and *ELSA Speak* provide adaptive interaction, instantaneous feedback, and psychologically safe environments for repeated speaking practice (Salsabil & Rakhmawati, 2025; Mardiah, & Saadillah, 2025; Huang, & Yang, 2023; Tai, 2022). These affordances renew a central question in applied linguistics: how does technology-mediated interaction influence both the affective and behavioral dimensions of second-language learning?

Speaking has long been recognized as the most anxiety-provoking aspect of EFL learning (MacIntyre, 2017). Fear of public mistakes and negative evaluation often constrains learners' willingness to communicate and undermines fluency (Brand & Götz, 2011). By providing private, judgment-free spaces for rehearsal, AI conversational agents can potentially reduce FLA and enhance self-confidence. Empirical studies from various Asian contexts indicate that sustained chatbot interaction lowers anxiety and increases self-efficacy in speaking tasks (Zhang, Meng, & Ma, 2024; Wang, Tao, & Cheng, 2024). Similarly, voice-based agents have been associated with measurable gains in fluency, lexical range, and pronunciation accuracy (Akhter, 2025; Xiao, Zhao, Sha, Yang & Warschauer, 2023).

Beyond affective relief, AI-mediated dialogue appears to encourage SRL—the learner's capacity to plan, monitor, and evaluate their own progress (Nhan, Hoà, & Quang, 2025). Platforms that deliver adaptive, reflective feedback can strengthen learners' autonomy and persistence (Zhang, 2025; Hashemifardnia & Kooti, 2025). Yet, some scholars warn that overreliance on automated correction may result in cognitive off-loading, where learners accept AI feedback passively without deep engagement (Trinovita, Nurchurifiani, Hastomo, Andewi, & Hasbi, 2025). This tension between support and dependency underscores the need for a nuanced understanding of how AI-mediated speaking practice shapes both emotional regulation and self-directed learning behaviors.

Despite growing interest, three critical gaps remain in current scholarship. First, most studies examine either affective or linguistic outcomes in isolation, offering limited insight into how emotional and cognitive mechanisms jointly contribute to oral proficiency. Second, comparative evidence across *instructional modes*—human-taught, AI-mediated, and blended—is scarce, especially in under-explored contexts such as North Africa. Third, few studies integrate complementary frameworks—*Positive Psychology* (Dewaele & MacIntyre, 2014), *Self-Regulated Learning theory* (Zimmerman, 2002), and the *Interaction Hypothesis* (Long, 1996)—to explain how affective, cognitive, and interactional processes converge in AI-enhanced oral communication.

Addressing these gaps, the present study investigates how AI conversational agents influence EFL learners' speaking anxiety, oral performance, and self-regulated learning within Algerian higher education. It also explores how learners perceive these experiences qualitatively through reflective journals, aiming to triangulate statistical and thematic evidence for greater explanatory depth. By situating AI use within an integrated affective-cognitive framework, the study contributes to both *CALL theory* and *pedagogical practice* in technology-supported speaking instruction.

Accordingly, the study was guided by the following overarching question: *How does the mode of instruction—AI-mediated, blended, or human-taught—affect learners' affective, behavioral, and self-regulatory dimensions in oral English communication?*

From this question, four focused sub-questions and corresponding hypotheses were derived:

- *RQ1:* To what extent does the mode of instruction influence foreign language anxiety (FLA)?
- *H1:* Learners in AI-mediated and blended groups will exhibit significantly lower FLA than those taught exclusively by humans.
- *RQ2:* To what extent does the mode of instruction affect oral performance, as measured by an analytic rubric adapted from the IELTS Speaking Band Descriptors?
- *H2:* AI-mediated and blended learners will achieve greater gains in fluency, accuracy, and pronunciation than the human-taught group.
- *RQ3:* To what extent does the mode of instruction affect self-regulated learning (SRL) behaviors during oral practice?
- *H3:* AI-mediated and blended learners will report higher levels of goal setting, monitoring, and reflection compared to their human-taught peers.
- *RQ4:* How do learners perceive and reflect on their affective and self-regulatory experiences in AI-mediated oral practice?
- *H4:* Qualitative reflections will reveal that AI-supported interaction reduces anxiety, increases motivation, and enhances metacognitive awareness, consistent with Positive Psychology and SRL principles.

By combining quantitative and qualitative analyses, this research aims to provide an empirically grounded and theoretically coherent account of how conversational AI reshapes the interplay between emotion, cognition, and performance in EFL speaking contexts.

Literature Review

The Dual Edges of AI Conversational Agents in EFL Speaking Development

The integration of AI into EFL speaking pedagogy represents a paradigmatic shift in how learners rehearse, monitor, and refine oral competence. AI conversational agents—ranging from rule-based bots to generative, voice-enabled systems—draw on natural-language processing and automatic speech recognition to deliver adaptive feedback and foster learner autonomy (Purwoko, Hidayati, & Rahmawati). Meta-analyses confirm significant gains in fluency, pronunciation, and overall communicative competence when AI tools are embedded within pedagogically structured tasks (Okyar, 2023; Pérez, Daradoumis, & Puig, 2020; Du & Daniel, 2024; Zheng, 2024).

Empirical evidence consistently shows that learners engaged in chatbot- or voice-based speaking outperform peers relying solely on traditional instruction (Vázquez-Cano, Mengual-Andrés, & Martínez, 2021; Yang, Kim, Lee, Shin, 2022). Applications such as ELSA Speak have improved both segmental and suprasegmental accuracy while enhancing confidence after brief interventions (Aulia, Sagala, & Ginting, 2025). These findings collectively portray AI as a productive complement—not a substitute—to human instruction by enabling individualized, data-driven feedback.

Yet the evidence also reveals limits. AI-mediated discourse frequently lacks the socio-pragmatic richness, turn-taking fluidity, and contextual nuance of authentic conversation (Lin, Chen, & Park, 2023; Park, 2024; Vanjani, Aiken, & Park, 2020). Technical errors, robotic intonation, or insensitive feedback can trigger frustration and demotivation (Çakmak, 2022). Some studies even note that perceived confidence gains may exceed measurable linguistic progress, suggesting an “illusion of mastery” (Wang, Tao, & Chen, 2024). These contradictions highlight that sustainable progress depends on teacher mediation and reflective debriefing that connect automated feedback to social interaction and pragmatic awareness (Du & Daniel, 2024).

AI and Foreign Language Anxiety

Within the Positive Psychology paradigm, affective variables such as enjoyment and anxiety fundamentally shape performance (Dewaele & MacIntyre, 2014). AI-mediated contexts may buffer anxiety by providing judgment-free rehearsal spaces and immediate, low-stakes feedback. Quantitative and qualitative studies generally corroborate this potential: sustained chatbot interaction tends to reduce FLA and enhance willingness to communicate (Zhang et al., 2024; Muthmainnah, 2024). For instance, Wang et al. (2024) found that embodied avatars with expressive faces fostered stronger anxiety reduction than text-only agents, while Hawanti & Zubaydulloevna (2023) reported that supportive AI feedback decreased writing anxiety and reinforced confidence.

Nevertheless, outcomes are not uniformly positive. Ballıdağ & Aydın (2025) and Çakmak (2022) observed that misinterpretations or impersonal feedback can heighten anxiety, especially among low-proficiency learners. Several investigations point to a novelty effect—initial excitement that fades without pedagogical scaffolding or teacher follow-up (Van Dijk, 2025; Nguyen, 2024). Hence, emotional benefits appear conditional on interface quality, task design, and the degree to which AI use is integrated with supportive human mediation. Sustained affective relief requires blended ecosystems where technology augments rather than replaces teacher presence.

Self-Regulated Learning and Autonomy in AI Environments

Self-Regulated Learning (SRL) theory (Zimmerman, 2002) provides a cognitive lens for understanding how learners manage engagement with AI feedback. Many AI platforms now include dashboards, adaptive prompts, and performance analytics that scaffold goal-setting, monitoring, and reflection (An et al., 2021). Empirical evidence suggests that chatbot-assisted practice can strengthen persistence and strategic repetition (Zhang, 2025; Hashemifardnia & Kooti, 2025).

However, scholars caution against automation bias—a tendency to accept AI judgments uncritically (Spatola, 2024; Trinovita et al., 2025). When learners treat AI feedback as authoritative rather than diagnostic, opportunities for metacognitive reflection diminish. Research indicates that AI tools effectively support the *forethought* and *performance* phases of SRL but contribute less to *self-reflection* unless guided by teacher intervention (Trinovita, Nurchurifiani, Hastomo, Andewi, & Hasbi, 2025). Thus, pedagogical scaffolding remains essential for internalizing metacognitive strategies, rather than externalizing control to algorithms.

SRL and FLA are also reciprocally related: stronger self-regulation enhances perceived control and lowers anxiety (Teng, 2023; Alvandi, Faruji, & Salehi, 2024), while reduced anxiety facilitates experimentation and self-monitoring. This reciprocity underscores the pedagogical value of integrated affective-cognitive frameworks, wherein AI systems encourage planning and feedback analysis while teachers foster reflective dialogue and social meaning-making.

AI Adoption and Contextual Constraints in Algerian EFL Settings

Despite international enthusiasm for AI-assisted learning, contextual barriers in Algeria and similar developing regions remain significant. Unequal access to stable internet, limited exposure to English-language technology interfaces, and varying degrees of digital literacy influence both adoption and outcomes (Hadef, 2025). Socio-economic disparities often determine device availability, while large class sizes and exam-oriented curricula restrict opportunities for interactive experimentation

(Moussaoui & Cheratti, 2024). Moreover, cultural attitudes toward automation—particularly the perception that "real learning" requires teacher presence—can affect learners' willingness to engage autonomously with AI tools (Leogrande, 2023).

These contextual factors suggest that successful AI integration demands institutional support, teacher training, and culturally responsive design that aligns with local pedagogical norms. In such settings, blended instruction, where AI practice supplements rather than substitutes classroom interaction, may represent the most viable model for sustainable implementation.

Overall, prior research affirms that AI conversational agents can enhance oral proficiency, reduce FLA, and strengthen SRL; however, findings remain fragmented, methodologically uneven, and geographically narrow. Most studies employ short interventions, rely on small convenience samples, and seldom examine affective and cognitive outcomes concurrently. Comparative evidence across AI-mediated, blended, and human-taught modes—particularly within North-African or Arabic-speaking contexts—is virtually absent.

The theoretical conversation likewise remains compartmentalized: Positive Psychology explains affective benefits, SRL accounts for metacognitive regulation, and the Interaction Hypothesis clarifies linguistic negotiation, yet few studies integrate these frameworks empirically. Addressing these deficits, the present study adopts a mixed-methods comparative design to explore how conversational AI influences speaking anxiety, oral performance, and self-regulation among Algerian EFL learners, thereby linking emotional resilience, cognitive autonomy, and communicative competence within a single explanatory model.

Theoretical Framework

Foreign-language learning extends beyond linguistic competence to include affective, cognitive, and social dimensions. The emergence of AI-driven conversational tools reframes these dimensions by mediating how learners experience, regulate, and enact communication. This study therefore integrates Positive Psychology (PP), Self-Regulated Learning, and the Interaction Hypothesis into one affective–cognitive–interactional framework for examining the effects of AI conversational agents on learners' FLA, oral performance, and self-regulation.

Positive Psychology and Foreign-Language Anxiety

PP (Seligman & Csikszentmihalyi, 2000) shifts the focus from deficit correction to human flourishing. In language education, PP highlights how positive affective states—enjoyment, resilience, and willingness to communicate—facilitate language growth (Dewaele & MacIntyre, 2014). Within this view, reducing FLA is essential to sustaining communicative engagement.

AI-mediated environments instantiate PP principles by offering low-pressure rehearsal spaces and adaptive, non-judgmental feedback that mitigate fear of negative evaluation (Muthmainnah, 2024). Studies consistently associate chatbot interaction with higher enjoyment and confidence (Zhang, Meng, & Ma, 2024; Wang et al., 2024). However, PP cautions that positive emotion must coexist with optimal challenge to prevent complacency (van Dijk, 2025). In this framework, PP predicts that emotionally supportive AI practice lowers anxiety and increases willingness to communicate—conceptually aligned with RQ1/H1 on FLA reduction.

Self-Regulated Learning Theory

Self-Regulated Learning Hypothesis (Zimmerman, 2002) explains the cognitive–behavioral mechanisms by which learners plan, monitor, and evaluate their progress. Its cyclical phases—forethought, performance, and reflection—capture the iterative nature of autonomous learning.

AI-mediated systems scaffold SRL by delivering instant diagnostics, visual progress analytics, and adaptive prompts that promote strategic monitoring (An et al., 2021). Evidence links such affordances to heightened persistence and metacognitive awareness (Zhang, 2025; Hashemifardnia & Kooti, 2025). However, excessive reliance on automation can cause cognitive off-loading, where learners accept feedback passively (Trinovita et al., 2025; Spatola, 2024). Hence, SRL theory clarifies how AI support must be pedagogically mediated to cultivate critical evaluation and self-reflection. Within this model, SRL predicts variations in self-regulatory behavior (RQ3/H3) and links affective control (PP) with linguistic outcomes.

The Interaction Hypothesis

The Interaction Hypothesis (Long, 1996) situates language development in communicative negotiation, where comprehensible input, output modification, and feedback drive acquisition. AI conversational agents simulate these mechanisms by providing adaptive reformulation, clarification requests, and repetition opportunities (Divekar et al., 2021; Tai, 2022). Human-like avatars sustain conversational flow and elicit engagement that enhances fluency and grammatical accuracy (Fathi, Rahimi, & Derakhshan, 2024; Wang et al., 2024).

Distinct from human interlocutors, AI agents allow learners to rehearse privately, pause, and retry—reducing performance pressure while maintaining feedback intensity (Salsabil & Rakhmawati, 2025). IH therefore accounts for the oral-performance dimension (RQ2/H2) and connects affective safety (PP) and self-monitoring (SRL) to observable communicative gains.

An Integrated Affective–Cognitive–Interactional Model

Synthesizing these perspectives yields an integrated framework in which:

- *Positive Psychology (PP)* provides the affective foundation: reduced anxiety and increased enjoyment create readiness to speak.
- *Self-Regulated Learning (SRL)* supplies the cognitive mechanism: learners plan, monitor, and reflect to sustain improvement.
- *Interaction Hypothesis (IH)* offers the linguistic mechanism: negotiation and corrective feedback translate effort into performance.

These dimensions are mutually reinforcing. Emotional security (PP) fosters deeper engagement with AI-mediated tasks (IH); cognitive control (SRL) ensures feedback is internalized; and repeated interaction (IH) consolidates both affective confidence and strategic awareness. Collectively, they explain how AI conversational agents can influence FLA, oral performance, and self-regulation simultaneously—addressing RQ1–RQ4 and H1–H4.

Methodology

Research Design

This study adopted a mixed-methods quasi-experimental design that integrated quantitative and qualitative data to provide a comprehensive understanding of the impact of instructional mode on learners' affective, cognitive, and linguistic outcomes. Quantitatively, a pre-test–post-test nonequivalent control group design was used to measure the effects of three instructional conditions—AI-mediated, blended, and human-taught—on FLA, oral performance, and SRL. Qualitatively, participants' reflective journals were analyzed thematically to document how learners experienced emotional, motivational, and metacognitive change during the intervention. A mixed-methods approach allowed triangulation of numerical patterns and personal experience, strengthening both internal inference and ecological validity (Creswell & Clark, 2018).

The study used a convenience sample of volunteers from an extracurricular language program; within that pool, participants were assigned to conditions using computerized random allocation. Because initial recruitment did not constitute a random sample of the wider student population, the design is reported as quasi-experimental with random assignment among consenting participants.

Participants and Sampling Procedure

A total of 150 third-year EFL students (68 males, 82 females; M = 21.4 years, SD = 1.1) from the Department of English at University of Batna 2, Algeria, voluntarily participated in this study. Recruitment was conducted via department announcements, e-mail invitations, and in-class briefings; inclusion criteria required participants to be third-year undergraduates, available for the full six-week schedule, and without diagnosed hearing or speech disabilities. Exclusion criteria included prior formal, intensive pronunciation training (more than six months) or current participation in external pronunciation programs. All participants were active members of an extracurricular language activity program, which provided a flexible setting for implementation outside regular coursework.

After consent, participants were assigned to conditions by computerized randomization (simple random allocation) implemented by a research assistant not involved in instruction or assessment; allocation lists were kept separate from instructional staff to minimize allocation bias. Because the initial sample was convenience-based rather than drawn randomly from the university population, the study is described as quasi-experimental with randomized allocation among volunteers.

To reduce potential confounding, baseline measures included prior English proficiency (departmental course average in English subjects), self-reported digital literacy (validated 6-item scale), and prior AI experience (frequency categories). These covariates—prior proficiency, digital literacy score, prior AI experience, gender, and age—were tested in preliminary models and included as control variables where they contributed to model fit or significantly predicted outcomes.

Pre-test comparisons of FLA, SRL, and oral performance revealed no statistically significant differences across groups ($p > .05$), supporting baseline comparability for primary outcomes. Summary statistics appear in Table 2. An a priori power analysis using G*Power 3.1 (Faul et al., 2009) indicated that a total sample of 150 (three groups, $\alpha = .05$) provided power ≥ .80 to detect medium effects ($f = .25$) in ANCOVA designs. A background questionnaire recorded that 113 participants regularly used AI tools, 30 occasionally, and 7 reported no prior AI exposure; prior AI experience was tested as a covariate and retained in sensitivity models.

Table 1. Baseline Equivalence Across Groups (Pre-test Means ± SD)

Variable	AI-mediated	Blended	Human-taught	$F(2,147)$	p
FLA	3.47 ± 0.42	3.45 ± 0.38	3.49 ± 0.40	0.12	.89
SRL	3.21 ± 0.37	3.24 ± 0.35	3.19 ± 0.39	0.24	.79
Oral Performance	2.63 ± 0.36	2.60 ± 0.35	2.64 ± 0.38	0.18	.84

Note. No significant baseline differences were found among the three instructional groups, indicating comparable starting levels before treatment.

Instruments

Foreign Language Classroom Anxiety Scale (FLCAS)

FLA was assessed using an adapted 20-item version of the Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986), rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). The instrument was reviewed by two bilingual experts for cultural appropriateness. Cronbach's alpha in the present sample was $\alpha = .91$.

Oral performance test

Participants completed pre- and post-intervention oral tasks rated with an analytic rubric adapted from IELTS Speaking Band Descriptors, covering fluency, pronunciation, grammar, and lexical resource. Two raters (PhD in Applied Linguistics, >5 years' assessment experience) completed calibration sessions using 20 pilot recordings. Raters were blinded to participants' group assignments and to the pre/post ordering of recordings; audio files were randomized and coded with anonymized IDs prior to scoring. Inter-rater reliability was ICC (2,1) = .87, 95% CI [.81, .91].

Self-Regulated Learning Questionnaire (OSLQ)

The Online Self-Regulated Learning Questionnaire (Barnard et al., 2008) was adapted to the oral-learning context, translated and back-translated, and piloted for clarity. Subscale reliabilities ranged from α = .79 to α = .88.

Reflective journals

Students submitted weekly reflective journals describing emotional responses, strategies, and perceptions of AI practice. A total of 732 entries (mean = 4.9 per participant) were collected and analyzed thematically. Table 3 summarizes instruments and reliability indices.

Table 2. Summary of Instruments and Reliability Indices

Instrument	Construct	Items	Scale	Reliability (α / ICC)
FLCAS	Foreign Language Anxiety	20	5-point Likert	α = .91
Oral Rubric	Speaking Performance	4	9-band	ICC(2,1) = .87
OSLQ	Self-Regulated Learning	24	5-point Likert	α = .79–.88

Note. All instruments demonstrated satisfactory to high reliability, ensuring dependable measurement across constructs.

Procedure

The intervention was conducted over six weeks during the first semester of 2024–2025. Instructional schedules were standardized across conditions so that total weekly speaking practice time was comparable. The AI-mediated group practiced with ELSA Speak (15 minutes/day, 5 days/week); the blended group combined daily 10-minute ELSA practice with a weekly 45-minute in-person session; and the human-taught group participated in a weekly 60-minute in-person oral expression class. *All in-person teaching was delivered by a single instructor who was also the researcher.*

Week 0 included an explanation of study procedures, written informed consent, and a technical induction to ELSA Speak for participants in the AI-involving conditions. Pre-test oral performance tasks were recorded under standardized conditions and anonymized prior to rating.

A shared task progression was used across groups to align learning objectives. The six-week sequence covered: Week 1—orientation and baseline tasks; Week 2—segmental accuracy; Week 3—word stress and rhythm; Week 4—connected speech and pausing; Week 5—fluency-building tasks; Week 6—integrated speaking and post-tests. AI-mediated participants completed corresponding ELSA modules curated to match each weekly target. In the blended condition, *the researcher–instructor incorporated learners' weekly ELSA error reports into planned activities,* using them to prioritize pronunciation targets and segmental or suprasegmental issues requiring clarification.

ELSA-based work consisted of segmental drills, prosody practice, shadowing tasks, and short monologue recordings. Classroom sessions (blended and human-taught groups) used instructor-guided pronunciation work, structured fluency practice, role-plays, and short communicative tasks aligned with the weekly focus.

Because the instructor was also the researcher, additional steps were taken to minimize expectancy and confirmation bias. These included:

- Using detailed lesson scripts specifying tasks, examples, and feedback phrasing to maintain standardization across sessions;
- Keeping the instructor blind to participants' pre-test scores during the intervention;
- Having an independent observer attend 20% of face-to-face sessions using a fidelity checklist;
- Ensuring that all pre- and post-test recordings were anonymized and rated by external assessors who were blind to group membership.

ELSA practice was tracked through application-generated logs (session length, modules completed) supplemented with weekly screenshots. Technical issues (e.g., login problems) were addressed individually, and any major disruptions were recorded for later sensitivity analyses.

Participants in the AI-mediated and blended groups submitted weekly reflective journals electronically. These journals documented learners' emotional responses, perceived progress, difficulties, and strategy use.

The researcher ensured compliance with approved ethical guidelines. *To reduce the influence of dual roles,* the researcher clarified verbally and in writing that participation, task completion, or performance would not influence coursework grades or evaluation within the department. All data were anymized and stored safely.

Data analysis

Quantitative analyses were performed in SPSS v29. Primary analyses for RQ1–RQ3 used ANCOVA on post-test scores with corresponding pre-test measures as covariates; additional control variables (prior proficiency, digital literacy score, prior AI experience, gender, age) were included where they improved model fit. Assumptions (normality via Shapiro–Wilk, homogeneity via Levene's test) were assessed; when assumptions were marginally violated, bias-corrected bootstrapping and heteroskedasticity-consistent standard errors (HC3) were used as robustness checks. Effect sizes are reported as partial η^2 and Hedges' g with 95% confidence intervals. Estimated marginal means were examined for interpretation.

Objective usage metrics from ELSA were used in two ways: (1) fidelity checks to confirm adherence to the prescribed practice schedule, and (2) as continuous predictors or moderators in secondary analyses to examine dose–response relations between usage and outcomes. Where usage varied substantially, per-protocol analyses were conducted alongside intent-to-treat comparisons.

Missing data were minimal (< 2% per variable). Primary analyses used complete-case data (N = 150); sensitivity analyses used multiple imputation (m = 20) under a missing-at-random assumption, and results were consistent across approaches. Influence diagnostics (Cook's D) identified three mildly influential observations; sensitivity analyses

Table 3. Overview of Experimental Learning Conditions

Group	Mode of Practice	Description
AI-mediated	ELSA Speak (15 min/day, 5 days/week)	Automated practice focusing on pronunciation, fluency, and intonation
Blended	10 min AI + 45 min face-to-face weekly	Combination of AI tasks and teacher-guided oral expression
Human-taught	60 min face-to-face weekly	Conventional oral expression with peer and teacher interaction

Note. Instructional conditions were designed to vary exposure to AI without altering task types, ensuring comparability across learning objectives.

excluding these cases produced the same substantive conclusions.

Qualitative analysis followed Braun and Clarke's (2006) thematic approach. All journals were imported into NVivo 12. Two coders independently read all entries and generated descriptive labels; through iterative meetings labels were grouped into themes. Intercoder reliability was estimated on a 20% subset (Cohen's $\kappa = .85$); disagreements were resolved through discussion and, when necessary, adjudicated by a third coder. Integration of quantitative and qualitative strands employed triangulation to align statistical patterns with narrative evidence.

Ethical considerations

All participants provided written informed consent after receiving study information on procedures, data use, and confidentiality. Participation was voluntary and withdrawal without penalty was assured. Use of ELSA logs required explicit opt-in consent; all usage data were anonymized and stored on secure university servers. Reported quotations from journals were anonymized and any potentially identifying details were removed.

Results

Quantitative analyses used ANCOVA models with the corresponding pre-test score as covariate to estimate adjusted post-test differences among three instructional groups (AI-mediated, blended, human-taught). The Results section is structured sequentially. First, we report the checks of statistical assumptions and robustness procedures. Second, we present descriptive statistics, estimated marginal means (EMMs), effect sizes, and Bonferroni-adjusted post-hoc contrasts. The section concludes by presenting the qualitative findings and their explicit integration with the quantitative outcomes, thereby addressing the fourth research question (RQ4).

Assumption Checks and Analytic Approach

Residuals from each ANCOVA approximated normality (Shapiro–Wilk on residuals): FLA $W = 0.988$, $p = .070$; Oral $W = 0.991$, $p = .142$; SRL $W = 0.989$, $p = .095$. Levene's tests for homogeneity of variance: FLA $F(2,147) = 0.92$, $p = .401$; Oral $F(2,147) = 3.12$, $p = .047$ (marginal); SRL $F(2,147) = 1.64$, $p = .199$. Given marginal heterogeneity for the oral outcome, oral-performance pairwise contrasts report bias-corrected bootstrapped 95% CIs (1,000 resamples). Tests of Homogeneity of Regression Slopes (Group × Pre-test) were non-significant for all models (FLA: $F(2,146) = 0.83$, $p = .437$; Oral: $F(2,146) = 1.12$, $p = .330$; SRL: $F(2,146) = 0.45$, $p = .639$), supporting ANCOVA assumptions. All models included the relevant pre-test score as covariate.

The analyses were conducted in SPSS v29 and R (version 4.3.2) using the emmeans and effect size packages. Model assumptions were verified through Shapiro–Wilk and Levene's tests, complemented by visual inspection of QQ and residual-versus-fitted plots. Influence diagnostics (Cook's D and leverage values) identified three mildly influential observations (Cook's $D > 4/n$). Sensitivity analyses excluding these observations and employing heteroskedasticity-consistent (HC3) standard errors produced the same pattern of significance and effect magnitudes, confirming the robustness of the results. Missing data were minimal (< 2% per variable) and handled through complete-case analysis ($N = 150$), with parallel multiple-imputation checks ($m = 20$) yielding equivalent outcomes.

Foreign-Language Anxiety (RQ1)

The descriptive (raw) pre/post means and SDs are presented in Table 1. After adjusting for baseline FLA, the instructional mode had a large, statistically significant effect on post-test FLA: $F(2,146) = 32.47$, $p < .001$, partial $\eta^2 = .31$. The pre-test covariate strongly predicted post-test FLA: $F(1,146) = 126.81$, $p < .001$, partial $\eta^2 = .46$.

Table 4. Raw Pre- and Post-Test Means (SD) by Group

Variable	Group	N	Pre-test M (SD)	Post-test M (SD)
FLA	AI-mediated	50	3.12 (0.57)	2.31 (0.48)
	Blended	50	3.11 (0.55)	2.65 (0.54)
	Human-taught	50	3.15 (0.60)	3.02 (0.61)
Oral performance	AI-mediated	50	6.15 (0.70)	7.88 (0.62)
	Blended	50	6.14 (0.69)	7.35 (0.68)
	Human-taught	50	6.12 (0.73)	6.92 (0.71)
SRL	AI-mediated	50	3.48 (0.50)	4.21 (0.43)
	Blended	50	3.46 (0.49)	3.92 (0.47)
	Human-taught	50	3.45 (0.52)	3.65 (0.52)

Raw means show that all groups began at similar baselines and diverged at post-test, with the AI-mediated group showing the greatest change across FLA, oral performance, and SRL. This table demonstrates baseline equivalence and justifies covariate adjustment in the ANCOVAs.

Estimated marginal means (adjusted for pre-test FLA): AI-mediated = 2.34 (SE = 0.05, 95% CI [2.24, 2.44]); Blended = 2.67 (SE = 0.05, 95% CI [2.57, 2.77]); Human-taught = 3.01 (SE = 0.06, 95% CI [2.89, 3.13]).

Bonferroni-adjusted pairwise contrasts (EMM differences):

- AI-mediated vs Human-taught: mean diff = −0.67, SE = 0.07, $p < .001$, 95% CI [−0.84, −0.50], Hedges' g = 1.24 (95% CI [0.92, 1.57]).
- Blended vs Human-taught: mean diff = −0.34, SE = 0.07, $p < .001$, 95% CI [−0.51, −0.17], Hedges' g = 0.63 (95% CI [0.34, 0.92]).
- AI-mediated vs Blended: mean diff = −0.33, SE = 0.06, $p = .012$, 95% CI [−0.57, −0.09], Hedges' g = 0.60 (95% CI [0.28, 0.92]).

These adjusted contrasts indicate large practical reductions in anxiety for AI-mediated learners relative to human-taught peers; blended instruction yields intermediate effects. The Hedges' g CIs exclude zero for primary contrasts, supporting the robustness of effects.

Oral Performance (RQ2)

ANCOVA controlling for pre-test oral performance showed a significant group effect on post-test oral scores: $F(2,146) = 23.68$, $p < .001$, partial $\eta^2 = .25$. The pre-test covariate was significant: $F(1,146) = 102.45$, $p < .001$, partial $\eta^2 = .41$.

EMMs (adjusted for pre-test oral performance; bootstrapped CIs used due to marginal heterogeneity):

- AI-mediated = 7.90 (SE = 0.09, bootstrapped 95% CI [7.72, 8.08])
- Blended = 7.37 (SE = 0.09, bootstrapped 95% CI [7.18, 7.56])
- Human-taught = 6.94 (SE = 0.10, bootstrapped 95% CI [6.75, 7.13])

Bonferroni-adjusted pairwise contrasts (bootstrapped):

- AI-mediated vs Human-taught: mean diff = 0.96, SE = 0.12, $p < .001$, 95% CI [0.68, 1.24], Hedges' g = 1.28 (95% CI [0.88, 1.67]).
- Blended vs Human-taught: mean diff = 0.43, SE = 0.11, $p < .001$, 95% CI [0.19, 0.67], Hedges' g = 0.57 (95% CI [0.30, 0.84]).
- AI-mediated vs Blended: mean diff = 0.53, SE = 0.11, $p = .018$, 95% CI [0.10, 0.96], Hedges' g = 0.72 (95% CI [0.33, 1.11]).

The AI-mediated group shows the largest adjusted gains in oral performance—consistent with a feedback-rich practice mechanism that accelerates the proceduralization of pronunciation and fluency. Bootstrapped CIs were used to address marginal heterogeneity and confirm the robustness of the primary contrasts.

Self-Regulated Learning (RQ3)

ANCOVA controlling for baseline SRL: $F(2,146) = 31.02$, $p < .001$, partial $\eta^2 = .30$. Pre-test SRL covariate: $F(1,146) = 89.72$, $p < .001$, partial $\eta^2 = .38$.

EMMs (adjusted for pre-test SRL):

- AI-mediated = 4.23 (SE = 0.05, 95% CI [4.13, 4.33])
- Blended = 3.93 (SE = 0.05, 95% CI [3.83, 4.03])
- Human-taught = 3.66 (SE = 0.06, 95% CI [3.54, 3.78])

Bonferroni-adjusted pairwise contrasts:

- AI-mediated vs Human-taught: mean diff = 0.57, SE = 0.07, $p < .001$, 95% CI [0.43, 0.71], Hedges' g = 1.12 (95% CI [0.80, 1.44]).
- Blended vs Human-taught: mean diff = 0.27, SE = 0.06, $p < .001$, 95% CI [0.15, 0.39], Hedges' g = 0.52 (95% CI [0.22, 0.82]).
- AI-mediated vs Blended: mean diff = 0.30, SE = 0.06, $p = .009$, 95% CI [0.08, 0.52], Hedges' g = 0.60 (95% CI [0.28, 0.92]).

AI-mediated learners reported greater increases in goal-setting, monitoring, and reflective study behaviors—processes that plausibly mediate the translation of frequent low-stakes practice into durable performance gains.

Table 5. ANCOVA Summary

Dependent variable	Source	SS	df	F	p	partial η^2
FLA	Group	16.28	2	32.47	< .001	.31
	Pre-FLA (cov)	63.91	1	126.81	< .001	.46
Oral	Group	22.33	2	23.68	< .001	.25
	Pre-Oral (cov)	96.22	1	102.45	< .001	.41
SRL	Group	12.87	2	31.02	< .001	.30
	Pre-SRL (cov)	34.20	1	89.72	< .001	.38

Table 5 reports effect sizes (partial η^2) and demonstrates that group explains a substantial portion of variance in each adjusted outcome.

Table 6. Estimated Marginal Means (EMMs) and SEs (adjusted)

Dependent	Group	EMM	SE	95% CI lower	95% CI upper
FLA	AI-mediated	2.34	0.05	2.24	2.44
	Blended	2.67	0.05	2.57	2.77
	Human-taught	3.01	0.06	2.89	3.13
Oral	AI-mediated	7.90	0.09	7.72	8.08
	Blended	7.37	0.09	7.18	7.56
	Human-taught	6.94	0.10	6.75	7.13
SRL	AI-mediated	4.23	0.05	4.13	4.33
	Blended	3.93	0.05	3.83	4.03
	Human-taught	3.66	0.06	3.54	3.78

Table 7. Selected Pairwise Contrasts (Bonferroni-adjusted)

Dependent	Contrast	Mean diff	SE	p	95% CI	Hedges' g (95% CI)
FLA	AI – Human	−0.67	0.07	< .001	[−0.84, −0.50]	1.24 [0.92, 1.57]
	Blend – Human	−0.34	0.07	< .001	[−0.51, −0.17]	0.63 [0.34, 0.92]
Oral	AI – Human	0.96	0.12	< .001	[0.68, 1.24]	1.28 [0.88, 1.67]
	Blend – Human	0.43	0.11	< .001	[0.19, 0.67]	0.57 [0.30, 0.84]
SRL	AI – Human	0.57	0.07	< .001	[0.43, 0.71]	1.12 [0.80, 1.44]
	Blend – Human	0.27	0.06	< .001	[0.15, 0.39]	0.52 [0.22, 0.82]

Qualitative findings (RQ4)

Qualitative analysis revealed four primary themes. They are represented in Table 8.

Table 8. Themes, Categories, Codes, and Illustrative Quotations

Theme	Category	Codes	Illustrative Extracts
1. Emotional regulation & confidence	Reduced anxiety	calm; less nervous; confident	"Before using the app I would freeze when I had to speak. After two weeks I could read aloud without panicking — the app let me try again privately, so I felt less judged and gradually more confident." (AI, F, 21)
	Positive emotional climate	comfort; security; enjoyment	"I actually looked forward to practice. The app's encouragement and the progress feedback made practice feel rewarding rather than stressful." (Blended, M, 23)
2. SRL growth	Goal setting	daily targets; self-monitoring	"I began to set small goals: three minutes on consonant clusters, then review. Tracking progress made me keep going even on busy nights." (AI, F, 22)
	Self-monitoring & feedback	reflection; awareness; correction	"When the app highlighted vowel errors, I recorded myself, compared, and corrected — I could see the exact change week to week." (Blended, F, 23)
3. Pronunciation & fluency development	Awareness of errors	articulation; segmental feedback	"Seeing the waveform and receiving segmental feedback helped me notice which sounds I was avoiding — I could target them directly in drills." (AI, M, 20)
	Fluency & automaticity	smooth speech; less hesitation	"At first I had long pauses; by week four I spoke more fluidly in the app and in short conversations with classmates." (Human, F, 24)
4. Technological & motivational factors	Engagement & motivation	enjoyment; gamification; persistence	"The progress bar and badges kept me practicing even when tired — I wanted to maintain the streak." (Blended, M, 23)
	Technical limitations	recognition errors; connection issues	"Sometimes the app didn't recognize me; it was frustrating but I learned to adjust microphone placement and still got useful feedback." (AI, F, 22)

The quotations illustrate how emotional safety and motivational affordances supported continued engagement with practice. Specifically, Theme 1 (emotional regulation and confidence) reduced affective barriers to speaking: learners reported feeling less judged, which increased their willingness to attempt speaking tasks and to repeat practice attempts until they felt improvement. This reduction in anxiety appears to have boosted both the frequency and the quality of practice, creating the conditions necessary for measurable gains in pronunciation and fluency. Theme 2 (SRL growth) describes the cognitive and metacognitive strategies—goal setting, self-monitoring, and targeted review—that participants adopted to structure practice; these strategies mediated the translation of increased practice into durable learning, and they align with the observed improvements on the SRL scale. Theme 3 (pronunciation and fluency development) captures the proximal, feedback-driven skill changes—greater segmental awareness and smoother delivery—that are the most direct contributors to measured oral-performance gains. Finally, Theme 4 (technological and motivational factors) explains how design features such as progress indicators and gamified elements sustained engagement over time and supported access to frequent, low-stakes practice; at the same time, occasional technical limitations moderated these benefits, highlighting implementation constraints to consider.

Taken together, these themes form an interlocking pathway in which emotional safety and motivating affordances increase engagement; SRL strategies determine how engagement is organized and optimized; and focused, feedback-rich practice produces proximal phonetic and fluency gains that register in objective oral-performance measures. In this way the qualitative evidence complements the quantitative results: reductions in foreign-language anxiety (FLA), increases in self-regulated learning (SRL), and improvements in oral performance reflect different points along a shared causal chain. Interactional Habits (willingness to initiate and sustain spontaneous speaking) emerge as a logical downstream outcome to test in future transfer assessments, given the combined influence of affective, metacognitive, and technological factors documented here.

AI-mediated instruction produced large, statistically significant improvements in foreign-language anxiety, oral performance, and self-regulated learning relative to blended and human-taught conditions (partial η^2 and Hedges' g reported above). Qualitative evidence further identified mechanisms—emotional safety, targeted feedback, goal-directed practice, and motivational affordances—that plausibly mediate these effects. Together, the quantitative and qualitative strands present a coherent, theory-driven account of how AI-mediated learning environments can enhance both affective and behavioral dimensions of language learning.

Discussion

This study investigated whether AI-mediated oral practice, compared with blended and traditional human-taught instruction, reduces FLA, enhances oral performance, and promotes SRL among adult learners following a six-week intervention. Across multiple outcome domains, AI-mediated instruction produced statistically robust and practically meaningful improvements. After controlling baseline scores, instructional mode significantly predicted post-test FLA, oral performance, and SRL, with large effect sizes. Pairwise contrasts revealed substantial standardized differences between the AI-mediated and human-taught conditions. These convergent quantitative results, supported by qualitative evidence of reduced fear, greater learner autonomy, and targeted feedback, suggest a dynamic interaction of affective, cognitive, and metacognitive factors.

The AI-based learning environment appeared to reduce affective barriers to speaking practice. Learners using AI reported markedly lower post-test anxiety than those in the human-taught group. Reflections from participants indicated that this effect stemmed from the tool's private, supportive nature and its capacity to provide immediate, individualized feedback. Lower anxiety likely freed attentional resources for pronunciation and fluency practice, which, in turn, contributed to the substantial improvements observed in oral performance.

The availability of the consistent, specific feedback together with opportunities for repeated practice also appeared to accelerate skill development. The AI group achieved the highest gains in oral performance and demonstrated stronger SRL behaviors than the other groups. This pattern supports a feedback-driven learning mechanism in which timely and precise feedback enhances error detection, correction, and strategic self-monitoring. Improvements in SRL suggest that AI tools not only facilitate technical practice but also strengthen learners' goal setting, planning, and evaluation skills, fostering more autonomous engagement with speaking activities.

The combined reduction in anxiety, growth in self-regulation, and improvement in oral performance supports an integrated explanation in which emotional comfort promotes engagement, engagement enables feedback-rich practice, and self-regulated strategies transform that practice into durable learning. This interplay of affective and cognitive mechanisms clarifies why AI-mediated instruction outperformed both blended and traditional modes of instruction.

The findings correspond closely with previous research demonstrating that conversational AI can enhance fluency, pronunciation, and affective comfort when embedded in structured language-learning tasks (e.g., Akhter, 2025; Vázquez-Cano et al., 2021; Zhang et al., 2024; Okyar, 2023; Muthmainnah). Studies employing applications similar to ELSA Speak have reported comparable improvements in pronunciation accuracy and learner confidence (Hoeriyah, 2024; Febrianti, 2025). However, not all evidence is uniformly positive. Some studies have identified mixed or minimal effects, emphasizing issues such as automation bias, limited socio-pragmatic authenticity, and short-term novelty effects (Çakmak, 2022; Lin et al., 2023; Wang et al., 2024; Van Dijk, 2025). These differences highlight that AI effectiveness depends heavily on design quality, pedagogical alignment, and the degree of teacher mediation.

Although the blended condition incorporated AI components, its outcomes were less pronounced than those of the AI-focused condition. A key explanation lies in how technology was integrated pedagogically. When AI is added as an auxiliary activity without redesigning tasks and feedback processes, its adaptive potential remains limited (Du & Daniel, 2024). Teachers in the blended condition may have treated the application as supplementary rather than central to instruction, thereby weakening the connection between automated feedback and classroom learning. Learners in the AI-mediated condition, by contrast, practiced more intensively within a consistent feedback environment, which likely contributed to greater fluency and confidence. The degree of teacher preparation and coordination in using AI also plays a role; effective integration requires deliberate planning to translate AI-generated feedback into communicative competence (Du & Daniel, 2024). In addition, the focus of measurement may have favored the AI group, as structured, task-specific assessments often reflect gains more directly associated with feedback-based practice than with spontaneous conversation (Lin et al., 2023; Park, 2024). Taken together, these explanations indicate that the success of AI-mediated instruction depends not merely on the technology itself but on how it is pedagogically embedded and consistently applied.

This interpretation helps reconcile divergent findings in the literature. Meta-analyses (Okyar, 2023; Pérez, Daradoumis, & Puig, 2020) report strong outcomes when AI tools are integrated into well-designed, feedback-oriented learning environments, whereas studies involving brief or loosely implemented interventions tend to show smaller or short-lived effects (Van Dijk, 2025; Nguyen, 2024). The current study therefore supports a conditional efficacy model: AI produces meaningful improvements when used systematically within a coherent pedagogical framework, but its benefits diminish when implementation is partial or inconsistent.

The findings contribute to theory in several ways. First, they identify AI-mediated practice as a potential mechanism for affective regulation in L2 learning, a dimension rarely emphasized in technology-enhanced instruction. Second, they extend socio-cognitive models of SRL by demonstrating that algorithmic feedback can act as an external regulatory cue that learners internalize to guide practice. Third, by showing consistent effects across affective, behavioral, and regulatory outcomes, the study clarifies the conditions under which adaptive feedback operates as an active pedagogical agent rather than a neutral medium. By integrating Positive Psychology (Dewaele & MacIntyre, 2014), SRL theory (Zimmerman, 2002), and the Interaction Hypothesis, the results propose an integrated framework in which AI lowers social-evaluative threat, promotes repeated corrective practice, and strengthens metacognitive planning and monitoring. These mechanisms collectively explain why previous studies have reported both strong and moderate effects depending on the degree of pedagogical coherence and contextual fit.

AI-mediated tools should be viewed as complementary rather than substitutive instructional resources. They are particularly effective for repetitive, low-stakes speaking practice that benefits from immediate corrective feedback. Successful use depends on aligning AI-based activities with classroom objectives and ensuring that teachers are prepared to interpret AI analytics and incorporate them into communicative lessons. Institutions should provide sustained professional development and technical support to ensure smooth integration.

Ethical and equity considerations are essential for responsible adoption. Concerns regarding algorithmic bias, data privacy, and cultural sensitivity must be addressed systematically (Spatola, 2024; Trinovita et al., 2025). Institutions should guarantee secure handling of learner data, evaluate system fairness, and prevent unequal access. These issues are particularly relevant in the Algerian context, where uneven connectivity and limited digital literacy may restrict engagement. Long-term success therefore requires infrastructural investment and culturally responsive adaptation that aligns with local pedagogical practices.

Several limitations should be acknowledged. The intervention spanned only six weeks, leaving the long-term stability of gains uncertain. Although the balanced sample enhances internal validity, participants represented a single educational context and proficiency level, limiting generalizability. Some measures relied partly on self-report, underscoring the need for triangulation with behavioral and acoustic data. Novelty and expectancy effects may have influenced results, and occasional recognition or connection issues highlight the importance of technical refinement. Furthermore, task-based assessment may not fully capture transferable communicative competence. Prior research suggesting a possible "illusion of mastery" (Wang et al., 2024) reinforces the need for longitudinal and ecological evaluations that assess both retention and authentic interaction.

Future research should examine long-term retention, spontaneous oral performance, and replication across diverse proficiency levels and educational contexts. Studies that isolate the effects of specific AI features—such as feedback precision, multimodal input, or motivational design—would clarify which aspects contribute most to learning. Further investigation into mediating and moderating variables, including anxiety level, motivation, and teacher engagement, could illuminate causal pathways. Analyses of cost-effectiveness and access equity are also needed to determine whether AI enhances or reinforces existing disparities. Rigorous, preregistered experimental and design-based approaches combining learning analytics with performance data will strengthen understanding of how AI supports sustained language development.

Conclusion

This study demonstrates that AI-mediated oral practice can substantially reduce foreign-language anxiety, increase self-regulated learning behaviors, and accelerate oral performance within a six-week intervention. The effects were statistically robust and practically meaningful, and qualitative themes converged on mechanisms of emotional safety, immediate feedback, and enhanced autonomy. While the findings advocate for the strategic inclusion of AI tools as supplements to classroom teaching, they also call for careful attention to long-term effects, contextual generalizability, and ethical considerations related to access and transparency. In sum, AI-mediated practice represents a promising, evidence-based addition to the language-learning toolkit — one that can amplify opportunities for low-stakes, feedback-rich speaking practice and, when paired with thoughtful pedagogical integration, help learners speak more confidently and effectively.

References

- Akhter, E. (2025). The impact of human–machine interaction on English pronunciation and fluency: Case studies using AI speech assistants. *Review of Applied Science and Technology*, *4*(1), 473–500. https://doi.org/10.63125/1wyj3p84
- Alvandi, M., Faruji, L. F., & Salehi, M. (2024). Relationship among EFL learners' self-regulated learning strategy use, speaking anxiety, and speaking strategy use. *International Journal of Instruction*, *18*(1), 379–396. https://doi.org/10.29333/iji.2025.18121a
- An, D., Zhu, C., & Wang, J. (2021). Examining self-regulated learning in digital environments: A systematic review. *Educational Research Review*, *34*, 100402. https://doi.org/10.1016/j.edurev.2021.100402
- An, Z., Wang, C., Li, S., Gan, Z., & Li, H. (2021). Technology-assisted self-regulated English language learning: Associations with English language self-efficacy, English enjoyment, and learning outcomes. *Frontiers in Psychology*, *11*, Article 558466. https://doi.org/10.3389/fpsyg.2020.558466
- Aulia, N., Sagala, R. W., & Ginting, P. (2025). The endorsement of self-regulated learning intercorporate with ELSA Speak AI to boost speaking skill of Thai EFL students. *Journal of English Language and Education*, *10*(2), 284–297. https://doi.org/10.31004/jele.v10i2.748
- Ballıdağ, M., & Aydın, S. (2025). A comparison of the effects of AI-based chatbots and peer interactions on speaking anxiety among EFL learners. *Future in Educational Research*, *3*, 224–238. https://doi.org/10.1002/fer3.70005
- Barnard, L., Paton, V. O., & Lan, W. Y. (2008). Online self-regulatory learning behaviors as a mediator in the relationship between online course perceptions with achievement. *International Review of Research in Open and Distance Learning*, *9*(2), 1–11.
- Brand, C., & Götz, S. (2011). Fluency versus accuracy in advanced spoken learner language: A multi-method approach. *International Journal of Corpus Linguistics*, *16*(2), 255–275. https://doi.org/10.1075/ijcl.16.2.05bra

- Çakmak, F. (2022). Chatbot–human interaction and its effects on EFL students' L2 speaking performance and anxiety. *Novitas-ROYAL (Research on Youth and Language)*, *16*(2), 113–131.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, *4*(2), 237–274. https://doi.org/10.14746/ssllt.2014.4.2.5
- Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning*, *35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, *6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, *121*, 103254. https://doi.org/10.1016/j.system.2024.103254
- Gusrianto, A. (2025). ELSA Speak-assisted pronunciation learning and its effect on EFL students' oral accuracy and confidence. *Journal of Applied Linguistics Research*, *8*(1), 22–37.
- Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: Alleviating students' anxiety in the English writing classroom. *Bulletin of Social Informatics Theory and Application*, *7*(2), 182–192.
- Hashemifardnia, A., & Kooti, M. (2025). AI-mediated language learning and EFL learners' self-confidence, self-regulation, well-being, and L2 motivation: A mixed-method study. *English Education Journal*, *16*(2), 109–124. https://doi.org/10.24815/eej.v16i2.45696
- Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.2307/327317
- Ji, H., Han, I., & Ko, Y. (2022). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*. https://doi.org/10.1080/15391523.2022.2142873
- Lai, C. (2024). Conversational AI in second language speaking: Pedagogical frameworks and empirical evidence. *Language Learning & Technology*, *28*(2), 55–77.
- Leogrande, A. (2023). Bridging the AI divide between developed and developing nations. *OSF Preprints*. https://doi.org/10.31219/osf.io/j9q84
- Lin, Y., Chen, H., & Park, M. (2023). Authenticity in AI-driven conversational learning: Challenges for pragmatic competence. *ReCALL*, *35*(3), 258–275.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie, & T. Bhatia (Eds.). *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J. Dewaele (Eds.). *New insights into language anxiety: Theory, research and educational implications* (pp. 11–32). Multilingual Matters.

- Mardiah, A., & Saadillah, S. (2025). Maximizing ELSA Speak for developing English fluency and reducing speaking barriers in language learners. *Issues in Applied Linguistics and Language Teaching*, *7*(1), 262–271. https://doi.org/10.37253/iallteach.v7i1.10421
- Moussaoui, W., & Cheratti, N. (2024). Analysis of the reality of digitalization and the digital divide in Algeria based on selected international indicators. *Dirasat wa Abhath: The Arabic Journal of Human and Social Sciences*, *16*(5), 498–513.
- Muthmainnah, N. (2024). Reducing foreign language anxiety through AI chatbot-mediated oral practice. *Asian EFL Journal*, *26*(2), 112–131.
- Nguyen, H. A. (2024). Harnessing AI-based tools for enhancing English speaking proficiency: Impacts, challenges, and long-term engagement. *International Journal of AI in Language Education*, *1*(2), 18–29. https://doi.org/10.54855/ijAILE.24122
- Nhan, L. K., Hoà, N. T., & Quang, L. V. N. (2025). Revolutionizing speaking skills improvement: AI's role in personalized language learning. *International Journal of Innovative Research and Scientific Studies*, *8*(2), 4637–4649. https://doi.org/10.53894/ijirss.v8i2.6408
- Okyar, H. (2023). Chatbots in English as a foreign or second language education contexts: A review of recent empirical research. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, *36*, 1333–1346. https://doi.org/10.29000/rumelide.1369170
- Park, M. (2024). Socio-pragmatic limitations in AI-mediated conversation for EFL learning. *Language and Technology*, *6*(1), 22–39.
- Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, *28*(6), 1549–1565. https://doi.org/10.1002/cae.22326
- Purwoko, D., Hidayati, A., & Rahmawati, N. (2023). AI-based conversational platforms and their effects on EFL oral proficiency. *International Journal of Language Education*, *7*(3), 41–56. https://doi.org/10.26858/ijole.v7i3.39881
- Safitri, E. I., Hidayati, S., & Ciptaningrum, D. S. (2025). The impact of AI chatbots on English language learners' speaking proficiency: A systematic review. *Journal of Research on English and Language Learning*, *6*(2), 317–329. https://doi.org/10.33474/j-reall.v6i2.23866
- Salsabil, A. D., & Rakhmawati, L. A. (2025). From silent learners to confident speakers: The effect of AI voice chat with ChatGPT on EFL speaking skills. *Journal of English Education*, *3*(1), 38–50. https://doi.org/10.61994/jee.v3i1.1137
- Seligman, M., & Csikszentmihalyi, M. (2000). Positive psychology. *American Psychologist*, *55*(1), 5–14. https://doi.org/10.1037/0003-066X.55.1.5
- Spatola, N. (2024). The efficiency–accountability tradeoff in AI integration: Effects on human performance and over-reliance. *Computers in Human Behavior: Artificial Humans*, *2*(2). https://doi.org/10.1016/j.chbah.2024.100099
- Tai, T. Y. (2022). Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom. *Computer Assisted Language Learning*, *35*(7), 1539–1563. https://doi.org/10.1080/09588221.2022.2075013

- Teng, X. (2023). Alleviating L2 writing anxiety and improving L2 writing performance through self-regulated learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4630683
- Trinovita, D., Nurchurifiani, E., Hastomo, T., Andewi, W., & Hasbi, M. (2025). Exploring the influence of generative AI on self-regulated learning: A mixed-methods study in the EFL context. *Jurnal Iqra': Kajian Ilmu Pendidikan*, *10*(2), 301–316. https://doi.org/10.25217/ji.v10i2.6389
- Van Dijk, W. (2025). Harnessing artificial intelligence to enhance speaking confidence in EFL learners with limited opportunities for real-world practice: A review and recommendations. *International Journal for Multidisciplinary Research*, *7*(2). https://doi.org/10.36948/ijfmr.2025.v07i02.42748
- Vázquez-Cano, E., Mengual-Andrés, S., & Martínez, J. (2021). Chatbots as language learning tools: A study on EFL speaking performance. *Education and Information Technologies*, *26*(4), 4875–4893. https://doi.org/10.1007/s10639-021-10552-0
- Wang, D., Tao, Y., & Chen, G. (2024). Artificial intelligence in classroom discourse: A systematic review of the past decade. *International Journal of Educational Research*, *123*, Article 102275. https://doi.org/10.1016/j.ijer.2023.102275
- Xiao, F., Zhao, P., Sha, H., Yang, D., & Warschauer, M. (2023). Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*, *4*(3), 300–325. https://doi.org/10.1515/jccall-2022-0032
- Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, *34*(3), 327–343. https://doi.org/10.1017/S0958344022000039
- Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System*, *121*, 103259. https://doi.org/10.1016/j.system.2024.103259
- Zhang, Z. (2025). The role of artificial intelligence tools on Chinese EFL learners' self-regulation, resilience, and autonomy. *European Journal of Education*, *60*(4). https://doi.org/10.1111/ejed.70127
- Zheng, S. (2024). The effects of chatbot use on foreign language reading anxiety and reading performance among Chinese secondary school students. *Computers and Education: Artificial Intelligence*, *7*, 100271. https://doi.org/10.1016/j.caeai.2024.100271

Appendices

Appendix A — Foreign Language Classroom Anxiety Scale (Adapted 20-item FLCAS)

Please indicate how much you agree with each statement below as it applied to your speaking practice during the study period. Use the following scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree.

Items

1. I feel nervous when I have to speak English in practice tasks.
2. I worry about making mistakes aloud when practicing English.
3. I get tense when I am asked to repeat spoken tasks.
4. I feel embarrassed speaking English in front of others.
5. I worry that my English pronunciation will sound bad.
6. I get anxious when I cannot find the right words while speaking.
7. I feel uncomfortable when my spoken English is being evaluated.
8. I feel my heart race when I must speak English under time pressure.
9. I feel confident speaking English in private practice (reverse scored).
10. I am afraid of asking for clarification when practicing speaking.
11. I worry others will judge my accent when I speak English.
12. I feel unsettled when a teacher or app highlights my speaking errors.
13. I often feel nervous before an oral task.
14. I feel comfortable trying new expressions when practicing speaking (reverse scored).
15. I feel pressured to perform well in speaking tasks.
16. I avoid speaking activities when I can (reverse scored).
17. I feel able to correct myself during speaking practice (reverse scored).
18. I worry that others will laugh at my spoken English.
19. I hesitate to start speaking in English because of fear of mistakes.
20. I enjoy speaking practice without worrying about being judged (reverse scored).

Scoring

- Reverse-score items: 9, 14, 16, 17, 20.
- Sum all items (after reversing) to create a total FLA score; higher scores indicate greater anxiety.
- Range: 20–100. For interpretation in group comparisons, present means (M) and SD and use ANCOVA with pre-test as covariate. Report Cronbach's alpha (α). In this study α = .91.

Administration notes:

Appendix B — Analytic Oral Performance Rubric (1–9 per dimension)

Scoring descriptors

Score	Fluency & Coherence	Pronunciation & Intelligibility	Grammar & Accuracy	Lexical Resource
9 (Excellent)	Speech flows naturally with few to no hesitations; discourse is cohesive and ideas connect logically.	Near-native segmental and suprasegmental control; highly intelligible without strain.	Accurate and complex grammatical structures used consistently; minimal to no errors.	Wide, precise vocabulary; idiomatic usage appropriate and varied.
7–8 (Very Good)	Generally smooth flow; minor hesitation that does not disrupt communication.	Clear pronunciation; occasional slips but fully intelligible.	Good control of grammar with occasional errors that do not impede meaning.	Good range of vocabulary; appropriate choice of words with some variety.
5–6 (Good / Competent)	Noticeable hesitations and self-corrections but overall coherent; occasional breakdowns in cohesion.	Generally intelligible though some segmental or prosodic errors affect ease of understanding.	Mix of simple and some complex forms; errors occur but meaning preserved.	Adequate vocabulary for most topics but limited precision and variety.
3–4 (Limited)	Frequent pauses, reformulations, and reduced coherence; message often interrupted.	Pronunciation problems frequently impede comprehension; listener effort required.	Frequent grammatical errors that sometimes obstruct meaning.	Limited vocabulary; repetitive and often imprecise word choice.
1–2 (Poor)	Speech is highly halting; ideas are fragmented; communication largely unsuccessful.	Severe pronunciation issues; largely unintelligible.	Persistent and severe grammatical errors that prevent comprehension.	Very limited lexical resource; cannot express basic meaning for many concepts.

Rater training and procedures

- Calibration: two calibration sessions using 20 pilot recordings; discuss anchor samples at scores 3, 5, 7, 9 for each dimension.
- Blinding: audio files anonymized and randomized; raters not provided group or time labels.
- Scoring: each rater scores all four dimensions independently. Final score per participant = average of raters' means across dimensions.
- Reliability: compute ICC(2,1) (two-way random, single measures) and report 95% CI. In this study ICC(2,1) = .87 (95% CI [.81, .91]). Report training materials and sample anchor recordings in Supplementary Materials if requested.

Administration of oral tasks

- **Tasks:** structured monologue (e.g., 1-minute description + 2-minute extended response) and semi-structured interview prompts. Use identical prompts at pre- and post-test with counterbalanced topic lists to avoid practice effects.

Appendix C — Online Self-Regulated Learning Questionnaire (OSLQ) — Adapted for Oral Practice

Scale and instructions to participants: Indicate agreement with each statement for the study period: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree.

Subscales and items

Subscales: (A) Environment Structuring (4 items); (B) Goal Setting (4 items); (C) Time Management (4 items); (D) Task Strategies (4 items); (E) Help Seeking (4 items); (F) Self-Evaluation (4 items).

Items

Environment Structuring

1. I arrange a quiet place for my speaking practice sessions.
2. I select times when interruptions are unlikely for my speaking practice.
3. I adjust my device or microphone to improve practice quality.
4. I organize my study materials specifically for oral practice.

Goal Setting

5. I set specific goals for each speaking practice session (e.g., target pronunciation of /θ/).
6. I establish short-term objectives (daily/weekly) for improving my speaking.
7. I monitor progress toward my speaking goals.
8. I revise my goals when progress stalls.

Time Management

9. I allocate time each day specifically for speaking practice.
10. I stick to a regular schedule for oral practice sessions.
11. I avoid procrastination when a speaking task is due.
12. I prioritize speaking practice over less-relevant activities.

Task Strategies

13. I use targeted drills to improve specific pronunciation features.
14. I record myself and compare recordings to track improvement.
15. I use AI feedback to create follow-up practice tasks.
16. I rehearse difficult parts repeatedly until they improve.

Help Seeking

17. I ask a teacher or peer for clarification when I do not understand feedback.
18. I consult online resources to resolve pronunciation questions.
19. I seek feedback from native speakers or more proficient peers when possible.
20. I request focused feedback on particular problems during lessons.

Self-Evaluation

21. I reflect on which practice strategies worked well for my speaking.
22. I note weaknesses in my speaking and make plans to address them.
23. I compare my current recordings with earlier ones to evaluate progress.
24. I adjust my study strategies based on self-assessment results.

Scoring and interpretation

- Compute subscale means (average of items in each subscale) and an overall SRL mean (average of all 24 items). Higher scores indicate stronger self-regulated practices. Report Cronbach's alpha for each subscale and the overall scale. In this study subscale α ranged .79–.88.

Administration notes: Administer pre/post; instruct respondents to answer with reference to their speaking practice during the intervention.

Appendix D — Reflective Journal Protocol and Coding Guide

1. Reflective journal protocol (instructions given to participants)

- Frequency & length: submit one journal entry per week for six weeks (~150–250 words per entry encouraged).
- Prompts: each journal should address the following prompts (respond in free narrative):
    1. Describe one or two speaking practice episodes from this week. What happened?
    2. How did you feel before, during, and after practice? (e.g., nervous, confident)
    3. What specific feedback did the AI or teacher provide, and how did you respond to it?
    4. What strategies did you use to improve your speaking (e.g., drills, recording, imitation)?
    5. What changes (if any) have you noticed in your speaking ability or confidence?
    6. Describe any technical or contextual issues that affected practice.

2. Submission logistics

- Entries submitted via secure university LMS or encrypted online form.
- Remind participants that entries are part of research (not graded) and that honesty is valued.

Саида Тоби

Универзитет Батна 2, Факултет за књижевност и стране језике,
Батна, Алжир

ВЕЖБЕ ГОВОРА ПОМОЋУ ВИ И ДИНАМИКА АНКСИОЗНОСТИ, УПОТРЕБЕ ЈЕЗИКА И САМОРЕГУЛАЦИЈЕ: ПОРЕЂЕЊЕ НАСТАВНИХ МОДАЛИТЕТА ПУТЕМ МЕШОВИТЕ МЕТОДЕ

У овом квазиекспериментaлном истраживању спроведеном коришћењем мешовите методе испитује се да ли вежбе говора помоћу ВИ (коришћењем ELSA Speak-а) смањују анксиозност у учењу страног језика (енг. Foreign language Anxiety – FLA), побољшавају усмено изражавање и подстичу саморегулисано учење (енг. Self-Regulated Learning – SRL) у поређењу са комбинованом и традиционалном наставом. У оквирима интегрисаног афективно-когнитивно-интеракцијског модела, који комбинује позитивну психологију, теорију саморегулисаног учења и хипотезу интеракције, уочене су мањкавости у упоредним доказима и истовременом мерењу афективних исхода и исхода у погледу употребе језика у учењу енглеског језика као страног у северноафричком контексту.

Сто педесет студената треће године основних студија енглеског језика као страног (n=150; 50 по групи) учествовало је у шестонедељном експерименту у којем је стандардизована прогресија задатака у различитим видовима учења (сегментни рад → прозодија → повезани говор → флуентност → интегрисани говор). Три вида учења била су: (1) учење помоћу вештачке интелигенције (ELSA Speak: 15 минута дневно, пет дана у недељи), (2) комбиновано учење (10 минута ELSA + 45 минута недељно уживо) и (3) учење уз помоћ наставника (60 минута недељно, уживо). Премере и постмере су укључивале прилагођени FLCAS од 20 ставки (α=0,91), аналитичку усмену рубрику (адаптирану из IELTS; ICC=0,87) оцењену од стране два спољна оцењивача и прилагођени онлајн SRL упитник (α=0,79–0,88). Дневници које су студенти писали сваке недеље (732 уноса) тематски су анализирани (NVivo; κ=0,85) како би се објаснили механизми. Квантитативни ефекти процењени су помоћу ANCOVA са коваријатима пре тестирања и проверама робусности (Bootstrapping, HC3 SEs, вишеструка импутација). Кориснички логови из ELSA Speak употребљени су за проверу аутентичности и секундарну анализу дозним одговором.

Резултати указују на велике предности учења посредовањем вештачке интелигенције. Након прилагођавања основних резултата тип наставе је значајно предвидео посттестни FLA ($F(2,146)=32,47$, $p<.001$, парцијално $\eta^2=.31$), ниво усменог излагања ($F(2,146)=23,68$, $p<.001$, парцијално $\eta^2=.25$) и SRL ($F(2,146)=31,02$, $p<.001$, парцијално $\eta^2=.30$). Парни контрасти (прилагођени Бонферонијевом методом) указали су на значајне разлике између учења помоћу вештачке интелигенције и традиционалног учења: средња разлика FLA=−0,67 (Хеџисов g≈1,24), средња разлика код усменог језика =+0,96 (g≈1,28), средња разлика SRL=+0,57 (g≈1,12). Показало се да комбинована настава има средњу, статистички значајну предност у односу на наставу коју држи наставник, али далеко мању од учења уз посредовање вештачке интелигенције. Анализе осетљивости (по протоколу, искључивање утицајних случајева) потврдиле су ове налазе.

Тематском анализом идентификована су четири међусобно повезана механизма који вероватно посредују у исходима: (1) емоционална регулација и самопоуздање (смањен страх од негативне евалуације и већа спремност за вежбање), (2) раст SRL-а (постављање циљева, праћење, рефлексија поткрепљена аналитиком апликације), (3) развој изговора и флуентности (циљане вежбе са много повратних информација и праћењем напретка) и (4) технолошке и мотивационе могућности (гејмифицирани индикатори напретка који одржавају истрајност; повремена ограничења у погледу препознавања/повезивања). Квалитативни подаци триангулисани су са квантитативним како би се подржао пут афективног→бихевиоралног→перформансног напретка: смањена анксиозност је омогућила чешће, фокусирано вежбање; вежбања су организована помоћу SRL стратегија; циљане повратне информације убрзале су напредак студената у погледу изговора и флуентности.

Краткорочне, интензивне усмене вежбе помоћу ВИ довеле су до значајног смањења FLA и приметног побољшања, како уочљивих усмених перформанси, тако и саморегулаторних понашања у поређењу са традиционалном наставом. Ови ефекти зависе од педагошког усклађивања: вештачка интелигенција функционише најефикасније када су циљеви вежбања кохерентни, повратне информације честе и интерпретабилне, а наставници интегришу аналитику у часове комуникације. Питања етичности и равноправности (приватност података, алгоритамска правичност, неједнак приступ апликацијама, дигитална писменост), ограничено трајање истраживања и узорак узет само са једног места указују на могућност генерализације. Стога је пожељно да се истраживање понови међуконтекстуално и на дужи рок.

Образовне институције би требало да размотре увођење ВИ алата као допуну усменим вежбама, јер је богата повратним информацијама и може да смањи афективне баријере и да негује SRL. Да би увођење ових алата било сврсисходно, потребно је (а) ускладити задатке генерисане вештачком интелигенцијом са циљевима учења, (б) обучити наставнике да тумаче и претварају аналитику вештачке интелигенције у комуникативне активности, (ц) планирати инфраструктуру и једнаке услове како би се избегло повећање јаза у приступу алатима и (д) спровести мере заштите података ученика и транспарентност система. Будућа истраживања требало би да тестирају дугорочну ретенцију знања, прелазак на спонтану интеракцију и специфичне карактеристике вештачке интелигенције (грануларност повратних информација, мултимодални сигнали, мотивациони дизајн) које покрећу одрживо учење.

Кључне речи: *настава помоћу ВИ, анксиозност у учењу страних језика, усмено излагање, саморегулисано учење, учење језика помоћу компјутерске технологије*