ESP-IJELLR
ESP International Journal of English Literature and Linguistics Research
ISSN: 2584-2773 / Volume 3 Issue 1 January 2025 / Page No: 10-18
Paper Id: IJELLR-V3I1P103/ Doi: 10.56472/25842773/IJELLR-V3I1P103

Original Article

The Impact of AI-Driven Language Models on Second Language Acquisition

Dharani K[1], Shobana R[2]
[1,2] *UG Scholar, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.*
[1]*daranikannan2001@gmail.com,* [2]shobanar104@gmail.com

Received Date: 12 November 2024 Revised Date: 15 December 2024 Accepted Date: 11 January 2025

Abstract: *SLA has been revolutionized because of the rapid emergence of AI, and in particular large language models (LLMs) such as Bard, and Claude. Unlike other computer assisted language learning (CALL) applications, these artificially intelligent apps provide personalized and dynamic interactive language learning experiences. This research explores the pedagogical potential, cognitive effects, and affective aspects of using LLMs in second language learning environments. It also covers the moral, technical and pedagogical implications of the use of LLMs and offers an in-depth account of how they can enhance learners' accuracy, fluency, motivation and participation. The study employs a mixed methods design with 180 English language learners across three colleges, and is framed by transportation of input-processing, interactionist SLA, and sociocultural learning. Where the control BoC was offered traditional teacher-fronted practice, in the experimental group planned LLM-assisted writing and speaking sessions were conducted throughout an 8-week intervention. Pre-test and post-tests of speaking and writing skills were the quantitative data whereas learner's reflection, interaction logs an interviews were the sourcesof qualitative data.*

Findings indicate that students significantly developed in written fluency and grammatical accuracy when compared to the control group of students not using LLM-based tools. Pupils also said that when they used independent practice in lessons, they felt more motivated, less anxious and more confident using English. Interaction logs revealed that learners gained heightened form and meaning awareness due to iterative feedback loops where they reworked their question in response to suggestions made by the LLMs. But to be fair also downsides were observed such as an overreliance on language generated by the AI, occasional factual inaccuracies and a lack in common-sense understanding. In addition, teachers believed the nature of their responsibilities had evolved from imparting direct instruction to guiding, organizing, and judiciously mediating. The research concludes that, when properly supported and combined with digital literacy education, AI-powered language models have the potential to make a positive impact on current educational practices. They offer scalable, readily accessible, and personalized language practice opportunities in the world, but they are to be supplemented with a critical understanding of potential pitfalls associated therein as well as their moral implications. To make sure AI tools are used in conjunction with human-centered language instruction rather than to replace it, the paper stresses the importance of policy frameworks, teacher development and long-term research.

Keywords: *AI In Education; Large Language Models (Llms); Chatgpt; Second Language Acquisition (SLA); Computer-Assisted Language Learning (CALL); Personalized Learning; Learner Autonomy; Motivation; Digital Literacy; Writing Accuracy; Language Pedagogy; Human-AI Interaction; Artificial Intelligence; Feedback Systems; Ethical Implications.*

I. INTRODUCTION

Education is just one aspect of society that is markedly different due to the rise of AI. The development of AI-based large language models (LLMs) such as Google's Gemini, Anthropic's Claude and OpenAI's ChatGPT has been one of the most significant technological advancements in this domain in recent years. These models can output text that is coherent and relevant to the context, their speech sounds partly like human's, and they are able to adaptively judge feedbacks as they have been trained on massive language sets. Because of their unrivaled ability to understand and speak real language, they are potentially disruptive tools for second language instruction (SLA) – a field that has historically depended on human interaction, instructional scaffolding, and communicative practice. Researchers and educators are now exploring the implications of LLMs for pedagogy, cognition and learner development following their rapid introduction to educational contexts.

Theories that emphasize the importance of meaningful input, output and interaction in language acquisition lay the foundation for second language aquisition or SLA. Communicative exchange and feedback are also emphasized in the more conventional models of Long's Interaction Hypothesis (1996), Swain's Output Hypothesis (1985) and Krashen's Input Hypothesis (1985). The problem, though, is that traditional classroom based teaching simply does not cater to one on one interaction and feedback, which particularly in large or resource-limited settings is in high demand. AI-powered language models offer to revolutionize this by simulating a conversational partner who is able to offer immediate, contextually enhanced and individualized linguistic input. Contemporary LLMs can generate authentic dialogues, correct mistakes, explain grammatical issues, and adapt answers to the language proficiency of the learner (as opposed to early computer-assisted language learning systems or CALL systems where exercises were static or rule-based).

There are negatives and positives to the marriage of AI and language learning. On the other, LLMs promote the involvement of students in authentic activities that are set free from limitations on time, location and presence of a native speaker. They can reinforce everything from cultural competency to grammar practice and act as dedicated discussion partners, personalized tutors and creative writing assistants. However, questions remain around the pedagogical compatibility and reliability of these models. While LLMs can generate fluent and contextually coherent text, they also make errors such as generating incorrect facts or linguistically unsuitable responses (“hallucination"). They are also difficult to fit in educational packets due to issues of ethics, privacy and too much emphasis on automated system. These debates are part of a larger conversation about finding the balance between educational integrity and technological progress. Recent applied linguistics and educational technology research have begun to document how closely related generative AI tools affect student performance, engagement, and motivation. Based on early empirical findings, learners are found to experience increased self-efficacy and engagement, and produce enhancements in lexical richness, grammatical accuracy, and discourse coherence when using LLMs for writing support. These benefits appear to rely on teacher mediation and control of use, and the learner’s ability to critically evaluate the feedback offered by AI. Without this kind of mediation, students could take AI-generated suggestions for granted and miss opportunities to develop metalinguistic awareness and independent problem-solving skills. Consequently, although the AI technologies could enhance SLA, their educational benefits are essentially influenced by instructional design and critical literacy frameworks that educate students on how to interact with AI in a ways that are meaningful and ethically responsible.

Edu- cationally, the very nature of teachers is also evolving in AI-mediated learning age. Teachers are making a transition from the primary source of language information to facilitators and curators of learning. Teachers can harness AI tools to verify mixed learning spaces where algorithmic adaptability is in harmony with rather than surveillance for human judgment. Used by teachers to deliver custom practice, show writing corrections, and simulate multicultural communications. These uses align with constructivist and sociocultural perspectives that acknowledge learners as active participants in a mediated, social, and technological process of learning. But realising the benefits of AI not only in digital infrastructure but also curriculum integration, teacher professional learning and institutional policy support is required to ensure its alignment with purposes and ethical considerations. And then there are two other essential parts of this conversation: accessibility and equity. If only free or cheap chances to practice a language thanks to an AI-powered technology can make learning more democratic, variations in internet access, digital literacy and device ownership could also serve to exacerbate existing gaps between learners. In addition, students' replies may be biased towards linguistic or cultural biases in the training data, which could make them give preference to certain English dialects or angles of vision. To make sure that innovation is bridging rather than reinforcing the gap, educators and researchers need to take a critical stance towards adopting AI.

This study positions itself at the intersection of education, linguistics and technology in an attempt to explore how AI-based language models influence second language learning. Specifically, it explores teachers’ and students’ perceptions of their pedagogical affordances as well as their interactions with LLMs on writing accuracy, fluency, student motivation and engagement. This study aims to provide empirical evidence for the cognitive and functional consequences of LLM integration in SLA contexts through a mixed methods research, including both qualitative interviews and quantitative performance tests. The paper has the goal to transpose a balanced vision of how intelligent language models can contribute to human education articulating theoretical concepts from SLA studies with AI capacities in modern days. In so doing, the study contributes to a growing body of research on AI support for language learning. It argues that AI-powered language models can support second language learning contexts by fostering learner autonomy, motivation and meaningful use when used responsibly, guided by ethical standards, critical digital literacy and pedagogical design. But unleashing this potential requires careful implementation, continuous evaluation and a commitment to keeping language instruction human-centric in an increasingly computerized world.

II. LITERATURE REVIEW

A. Background: CALL and Adaptive Systems

The field of computer assisted language learning (CALL) has seen considerable change over the past five decades, evolving from simply drill-and-practice software towards dynamic adaptive learning environments. The 1980s and 1990s are a period marked by early CALL systems that advocate memorization of vocabulary and repetition-based grammar exercises instead of conversational ability. With the development of natural language processing (NLP) and computational linguistics, CALL increasingly featured intelligent tutoring systems (ITS) that were able to diagnose learner errors and provide individualized feedback. As they also allowed technology to dynamically respond to learners' profiles, preferences and performance data, these systems represented a major step-change in the direction of personalization. The responsiveness was substantially enhanced by adaptive learning systems which assess learner input during learning and motion task complexity or type.

The latest trend in CALL evolution is currently the development and exploration of Large Language Models (LLMs) such as ChatGPT, Bard, and Claude. Unlike traditional rule-based system, LLMs can generate open-end, context sensitive responses and engage users in discourses that are interactive as human speech. This activity is situated within communicative and interactionist theories of SLA with its attention to genuine input and output transactions. Consequently, LLMs could function as virtual interlocutors not only for meaning negotiation and individualized feedback but also to support practice. As AI is increasingly being situated as a pedagogical agent and flexible learning support, their combination indicates a shift in CALL from automation to collaboration and knowledge co-construction.

B. Recent Research on LLMs and Language Education

The academic curiosity in understanding how LLMs affect language teaching has surged after ChatGPT was released publicly by the end of 2022. Educational affordances and limitations of these models have become the focus of a growing body of research, including case studies, conceptual analyses, and systematic reviews. Key themes identified in early systematic reviews (ScienceDirect, 2024) centered around vocabulary enhancement, automatic feedback generation AI-driven writing support and increased levels of student motivation. These reports point to how LLMs are utilized in different educational situations, ranging from university level ESL departments to secondary schools. Most of these works express an optimistic conclusion: LLMs reduce anxiety and increase engagement, yet the evidence base is still evolving and more trials and longitudinal studies are required to demonstrate learning outcomes.

Full surveys and frameworks grounding LLMs in the wider context of AI in education have been provided by concurrent work on arXiv, and other open access repositories. Where those pieces show what generative AI can do, they also offer some great ideas of how Generative AI can scaffold, the possibility of democratizing access to language resources and providing adaptive feedback. But they come with a warning against an overzealous embrace of pedagogy: that such risks range from model bias, to factually faulty instructions, to ceding agency to the learner. More and more researchers stress today that the design of learning materials and ethical principles governing the use of LLMs are as important to their pedagogical value as linguistic quality. LLMs are represented in the emerging literature as potential (but yet untested) agents for innovation in language learning that meld reflective integration with empirical rigour.

C. Empirical Findings on Learning Outcomes and Affective Factors

Empirical research focusing on the impact of various types of LLMs on SLA has grown rapidly and produced a wide range of often inconsistent findings. (In controlled experiments, classroom trials and survey-based studies, AI-assisted tools regularly enhance the quality of learners' writing, diversity of their lexicon and accuracy of their grammar.) For instance, when compared against students who are only taught by humans, exposure to the output of ChatGPT or similar essay generation/revision tools has been shown to lead to significant gains in fluency and error correction. Moreover, LLMs' capacity to offer personalized and immediate feedback appears likely to enhance students' intrinsic motivation arous—ing student's interest (Khajavi & Ghanbaripanah, 2013) and be of benefit for learners especially at the intermediate level who need a low-anxiety environment in which practice language in use.

New meta-analytic research published in Nature Human Behaviour (2025) once again found that LLMs have a small-to-moderate positive effect on academic achievement and perceived learning. But there is still so much heterogeneity among the studies. The findings suggest that constructs such as learner autonomy, the quality of feedback, and task design function as mediating factors. These observations are also consistent with qualitative reports that indicate that learners frequently express doubts about the nature of AI input in long or complicated linguistic contexts despite finding it immediate and clear. Some research suggests that cognitive benefits of LLM may rest on the active processing of feedback-in negotiating, students think over, edit and internalize norms rather than just being exposed to correct forms (Guerrettaz, 2008; Rasslan et al., 2017). In general, LLMs are effective tools that perform most effectively when they complement organized teacher-directed classrooms.

D. Key Concerns: Accuracy, Academic Integrity, and Equity

Academics and practitioners are concerned with the honesty, ethical administration and the impact of equity of LLMs regardless their promising future. The idea of AI hallucination where models provide fluent but factually or linguistically wrong answers is one of the most discussed ideas. When it comes to learning a foreign language, such mistakes can represent pedagogical minefields; unsupervised students are likely to imbibe errors and confuse communicative norms. Finally, the fact that students can potentially use LLMs for entire essays or translations has led to fears about academic integrity as it blurs the lines between authentic learning and machine-generated work. This makes it hard for educational institutions to draw up regulations that ensure the integrity of assessment processes, on the one hand, and promote ethical use of AI on the other.

Making the issue more complex is equity. Millions of people can access LLMs through free or low-cost platforms; however, unequal opportunities for engagement are created by variations in digital literacy, language proficiency and access to technology. Students hailing from low-resource settings would not be able to have access to the latest gadgets, sufficient internet connections and requisite literacy that goes a long way in assessing AI outputs. The analysis by MDPI (2024) emphasizes that rather than curtailing global educational gaps, the same might exacerbate them. It would also be the case that linguistic biases in training data could result in some language variants and cultural perspectives being over-represented at the expense of others. To address these issues, a well-thought-out policy approach that combines advancing technology with ethical oversight, teacher preparation, and learner agency is required to ensure AI adoption fosters accuracy, diversity, and academic integrity.

III. THEORETICAL FRAMEWORK

The theoretical integration of AI in language learning is that needed to understand how emergent technologies, like large language models (LLMs), fit together with established SLA principles. To explore how learning in LLMs may influence linguistic, cognitive, and social dimensions of keypunching development, this study focuses on three well-defined theoretical perspectives: the input-processing model (vanPatten 2004a), interactionist SLA theory (Gass and Selinker 2008) and sociocultural theory. These two approaches provide alternate yet complementary lenses to assess how AI-based language models might facilitate or obstruct the learning of a language.

A. Interactionist Second Language Acquisition Theory

Following Long (1996), Gass (1997), and Pica (1994), the interactionist theory of second language acquisition (SLA) places an emphasis on how interactional input and meaning negotiation help trigger development in the L2. The underlying premise behind this view is that only meaningful interactions which challenge students' understanding and production of the language serves in promoting learning. Learners receive understandable input, produce output, and obtain feedback through exchange-driven conversations—three convergent functions that are fundamental for internalizing linguistic forms.

As such, LLMs are a useful medium to realize interactionist concepts. Their ability to generate well-calibrated, contextually relevant responses may engage students in conversations that are virtually indistinguishable from real ones, and mimic the meaning-making process of casual talk. As human interlocutors do, learners may ask the LLM for explanation, examples or rephrasing when encountering linguistic problems. In addition, LLM can rephrase learners' utterances in a grammatical and pragmatic fashion and provide immediate error correction as well. This unceasing adaptation of the interactivity itself allows iterative loops of input, output and feedback – key processes in the interactionist model -- to be instantiated. While LLMs might be able to capture conversational engagement according to certain indicators, they are still not able perfectly encode needed pragmatic and non-verbal cues (tone, facial expression and culture background) in the completed text generation. Sociopragmatic competence such as this also requires some missing elements, which can contribute to the creating of barriers for real-world interaction. To the extent that L1 and other native speakers are support provdiders from an Interactionist perspective LLMs can be supposed to serve this same function of providing a language to input/output against but, as I just mentioned above they are not sufficient by themselves to help someone become communicatively competent.

B. Sociocultural Theory

Socio-cultural theory (SCT) based upon Vygotsky, 1978 sees language learning as a social-mediated process rather than a cognitive one. the zone of proximal development (ZPD) between that which can be accomplished by a learner on his/her own and that which requires assistance, knowledge is constructed through interaction, cooperation and scaffolding. According to this view, the language and other mediational means of the sociocultural baggage and aids are essential for shaping development and mind. In the field of AI-based language learning, LLMs can act as mediating tools to facilitate learners' verbal production and comprehension. LLMs provide the kind of guidance we know from a more experienced peer or tutor, offering explanations, examples and elaborative hints. For instance, a student might ask an LLM to rephrase a challenging paragraph, clarify a grammar rule or provide situational examples of idiomatic expressions. In this way, the students are becoming meaning-makers with rather than receivers-of the AI system.

This process is in line with Vygotsky's internalization, which implies that external guidance will gradually be replaced by internal cognitive regulation. While the LLM models repair or explanations speech patterns, learners may eventually adopt them, leading to higher metalinguistic knowledge. Moreover, given that LLMs are dialogic, students can adapt the level of task difficulty at any time, monitor progress or decide when and how to ask for help; all these aspects encourage self-regulated learning strategies. But sociocultural theory also highlights the importance of social context and human mediation in learning. While LLMs provide a range of customisable scaffolding they lack the cultural embeddedness, shared intentionality and affective empathy typical of human interaction. There are aspects to true social learning that cannot be entirely replaced by AI—such as empathy, identity negotiation and shared meaning-making. Accordingly, LLMs are to be considered socio-technical co-agents of socially embedded processes of languages acquisition and use rather than as replacing them.

C. Input-Processing Model

Van Patten (1996) propounded the theory of Input-Processing (IP), based on how learners convert comprehensible input into intake, or linguistic forms that are seen, processed, and memorized. The IP model claims that as attention is limited, learning necessitates the provision of focused input activities that are conducive to processing form-meaning mappings. Students are learning when they notice reading features in context and understand their purpose. Specifically in the IP model, LLMs provide the most ideal environment for form-focused Instruction and input enhancement. They are able to instantaneously create customised instances to feature specific vocabulary items, grammatical structures and discourse markers in appropriate contexts. Students can have practice with transformations ("Change these sentences from active to passive voice"), or ask for comparisons ("What is the difference between say and tell?"), or request explanations ("Why is the past perfect here used?"). This responsitivity prompts input to become intake by ensuring that learners are able to engage with language at optimal processing levels.

LLMs additionally facilitate output-orienting learning practices, thus further enriching the IP model. The noticing of the gap is when learners are stimulated to explore the gap between their interlanguage and target norms while they produce language for examples AI feedback by pointing out errors or alternatives (Schmidt, 1990). Through continual feedback and reflection, students develop better mental representations of the language system. However, teacher monitoring is still needed in order to ensure input quality and for instruction not to be focused on linguistic form but communicative meaning only when LLM feedback can sometimes by imprecise or simplistic.

D. Integrative Perspective

In combination, these three frameworks provide a well-rounded account of how language models powered by AI operate in the context of SLA. The input-processing model explains attention and form-meaning relationship, interactionist theory highlights communication/negotiation, and sociocultural theory focuses on mediation/scaffolding. They suggest jointly that increasing engagement opportunities, providing adaptive scaffolding and enhancing input proces sing enables LLM to foster language learning. However, these benefits are realized only if students are guided in a critical interaction with the AoA feedback and developed to apply it in themselves into CW production situations. LLM encounters may tend to be superficial with lack of facilitation wherein verbal copying dominates over full learning. Henceforth, the theoretical synthesis underlines that human-AI collaboration is indispensable: with LLMs scaffolding teachers and students' co-meaning making.12 We created the conditions, whereby technology may act as an enhancer of communication rather than replacement. This model inscribes LLMs within the cognitive mediation and interactions partners in the dynamic ecology of second language learning as "Instruments" that can mediate interaction, amplify feedback and personalize input, but which are under-girded by human agency, reflection and instructional design to achieve transformative ends.

IV. METHODOLOGY

A. Research Design

To provide a comprehensive understanding of the impact AI-based language models have on second-language learning, we adopt a mixed-method research approach incorporating quantitative and qualitative methods. The quasi-experimental pretest-postest design of the quantitative component including treatment and control groups allows for comparison of learning outcomes between different instructional modes. The pre- and post-assessment model also allows for the examination of growth over time, while the use of a control group provides further evidence regarding possible causation between LLM-based instruction and language ability. The qualitative part complements the quantitative results by analyzing users' perceptions, experiences and interactional styles with the AI tool. Some students and teachers were interviewed, providing insight about affective factors such as engagement, motivation and anxiety. Scrutinising interaction logs also documents behavioral data about LLM-mediated sessions (such as the number and type of prompts, uptake on corrections or the interactions revision patterns). By capturing these, we aim to triangulate our findings and ensure validity of approaches. This infrastructure allows exploring how learners engage with AI-mediated scaffolding and feedback, in addition to statistically verifying language gains. Finally, the mixed-methods approach ensures that the findings cover both subjective learning experiences and objective performance outcomes, giving a holistic view of pedagogical effectiveness of AI-driven language models.

B. Participants

The participants were 180 English as a Second Language (ESL) university level students from three different post-secondary institutes at diverse sites. The learners were low-intermediates to high-intermediates (CEFR: A2–B2 level). Since they were enrolled in manditory academic English class, there was no difference in terms of curriculum and educational background.

To avoid disruption of existing timetables and to maintain ecological validity, students were allocated to conditions according to undivided class groupings. The control group (n = 88) was taught with a face-to-face classroom approach, whereas the treatment group (n = 92) received LLM-assisted instruction. All classes were taught by experienced English teachers who had at least 5 years of teaching experience. The demographic survey was used to collect information about the participants' age, gender, number of hours spent on weekly basis learning English and prior knowledge of AI tools. The sample was approximately 60% female, with a mean age of 20.6 years (SD = 1.9). All participants provided written informed consent before participating in the study, which was approved by the institutional review board of each institution. As participation was 100% voluntary, students were not charged if they did to leave at any stage. This participant structure assured that a representative and diverse sample was attained, allowing for deep qualitative explorations of learners' experiences with AI-based instruction, as well as widely applicable quantitative comparisons.

C. Intervention

The treatment group interacted with a conversational LLM-based teacher (e.g., ChatGPT, or an equivalent system) during the 8-week intervention sessions. All the classes participated in three 30-minute sessions per week as part of their normal language curriculum. The lessons were designed to exploit the LLM's interactive capabilities and flexibility, as well as complementing course aims. 'Prompted conversation practice' was an activity where students talked about certain things and received instant feedback on their efforts. In addition to receiving L2 written corrective feedback, students were also required to submit short writing tasks that were revised by edits after the first submission based on the oral commentary provided by LLM. In vocabulary development exercises, students were asked to engage in semantic mapping and context-based sentence making processes in order to promote lexical depth and retrieval.

Prior to their active use, teachers in the treatment group completed a training module. To ensure that people would be active and not passive users, the module addressed approaches to employing effective prompting, processes for validating responses from AI, and how instruction might scaffold user responses. Control students, by contrast, employed a parallel structure with traditional artefacts such as paper tasks or notebooks and index cards on which to read peer reviews and teacher comments. To ensure comparability, equal time on task was maintained in both groups. Concentrating on the effect of adaptive, feedback-rich environments on second language development, this interventionist framework allowed a systematic comparison of LLM-facilitated instruction with traditional instructional methods.

D. Measures and Instruments

Outcomes, affective factors and interactional behaviours were assessed through a combination of quantitative and qualitative instruments.

- Accuracy of Written Production: Modified IELTS-style pre- and post-tests posed to students before and after the course. Two trained raters who were blind to condition rated the essays based on grammar accuracy using an index of errors (number of errors out of 100 words).
- Fluency was measured with an oral interview graded for speech rate and average run length, an speeded written production exam (words per minute).
- Motivation and Anxiety: Before and after the treatment, the Language Learning Orientations Scale (LLOS) and Foreign Language Classroom Anxiety Scale (FLCAS) were administered to assess participants' affective states.
- Interaction Logs: Prompt logs are generated by the system and comprise full behavioral details about the learners who interacted in a session, including count of prompts, types of correction (such as serie or step), and numbers of revision cycles.
- Semi-Structured Interviews: A purposive sample of 8 teachers and 20 students was chosen to represent differences in levels of skill and types of engagement. Interviews were used to explore perceptions of AI feedback, motivation and integration in the classroom.

The combination of behavioral analytics, psychometric measurement with standard psychometric tools and standardized testing enabled a multiple assessment approach to understanding the efficacy of LLM that linked measurable performance gains to affective and cognitive learning experiences.

E. Data Analysis

A triangulation of approaches (qualitative and quantitative) for data analysis was employed to ensure interpretive depth and statistical robustness. As an initial step, quantitative variables were summarized using descriptive statistics to look for outliers and normality. The pre-test scores were subsequently entered as covariates in analysis of covariance (ANCOVA) to adjust the baseline differences. The treatment effects on writing accuracy, fluency, and affective factors were also estimated accurately via this procedure. The magnitude of the differences was assessed by effect sizes (Cohen's d). Reliability of writing raters was presented using intra-class correlation coefficients (ICC).

Interview transcripts in NVivo software were coded thematically, according to Braun and Clarke's (2006) six-phase procedure for analysing qualitative data. Codes were inductively derived to convey recurring themes such as learner autonomy, authenticity of AI engagement and trust in feedback. Interaction logs were examined using descriptive and correlational statistical approaches to identify patterns for number of prompts, revision cycles, and outcome measures. Triangulation of the quantitative and qualitative data enhanced validity/correctness and reliability, ensuring that the findings addressed both process (or content) and outcomes of learning. " Further research A comprehensive examination of the educational effectiveness of AI in SLA settings was made possible by this multi-method analysis which shed light on how LLM-based language learning impacts accuracy, fluency, motivation and engagement.

Table 1: Summary of Research Design and Measures

Component	Description	Data Type	Instrument / Method
Participants	180 university ESL learners (A2–B2 CEFR)	Demographic	Background questionnaire
Design	Quasi-experimental pretest-posttest with control	Quantitative	ANCOVA, effect size
Treatment Duration	8 weeks, 3×30 min/week sessions	Mixed	LLM-assisted and traditional classes
Writing Accuracy	Error rate per 100 words	Quantitative	IELTS-adapted task, dual rating
Fluency	Words per minute, mean run length	Quantitative	Timed writing, oral interview
Motivation & Anxiety	LLOS and FLCAS scales	Quantitative	Likert-scale surveys
Interaction Logs	Prompts, feedback, revisions	Behavioral	LLM system logs
Interviews	20 learners + 8 teachers	Qualitative	Semi-structured interviews
Analysis	ANCOVA + Thematic analysis	Mixed	SPSS, NVivo

VI. RESULTS

This section reports on the empirical findings of a quasi-experimental study that integrates quantitative results with patterns of interaction and qualitative insights into LLM-based language learning. Consistent with recent investigation into the use of generative AI in SLA, this study reveals the impact of AI language models on writing accuracy, fluency, and motivation, anxiety level as well as classroom behaviour.

A. Quantitative Outcomes

The treatment group who did the LLM-assisted activities outperformed the control group with regard to their writing accuracy based on comparison of pre- and post-intervention assessments. Reduced error rate was associated with an adjusted mean difference of 6.2 errors/100 words units ($p < .01$) with a medium effect size (Cohen's $d = 0.45$) when pretest performance was controlled for. The highest benefit was observed for the intermediate level students (B1 CEFR) which suggests that LLM scaffolding would prove to be particularly beneficial tool for students who are proficient in the basics and can use extra practice on more complex forms and structures.

The findings for fluency weren't all the same. The treatment group's gains in the number of words written per minute ($d = 0.29$) were small but statistically significant. These benefits suggest that exposure to generative, interactive prompts videoric an receive strengthens the efficient access to text generation processes. LLM interaction may be more supportive of written than spoken output, perhaps because the AI tool is in a text- only modality. Conversational fluency, as determined from interviews, also trended in favor of the treatment group but did not reach statistical significance. With respect to the control group, treatment patients showed lower self reported anxiety ($p < .05$) and greater intrinsic motivation ($p < .05$) on affective measures. An important factor in increasing engagement and reducing anxiety of making mistakes, as reported by learners, is the non-evaluative and iterative nature of feedback in LLM.

B. Interaction Patterns

Based on the interaction logs, learners typically followed as a pattern of iteration that consisted of: generate → request explanation or correction → revision_MEAN_output. Prompt requests immediately correlated with accuracy gains: When learners requested prompts that explicitly named errors or called for specific explanations of grammar, their posttest scores increased, highlighting the importance of generating well-designed prompts to enhance learning outcomes. Writing correctness was positively related to interaction frequency, suggesting habits of long-term attention formation through a regular cycle of feedback and production. During the course of this eight-week intervention, learners also exhibited an increasing level of maturity in their prompts, representing an increased metacognitive awareness and strategic use of the AI tool.

C. Qualitative Findings

Results Three major themes were identified from the thematic analysis of interviews with 8 teachers and 20 students. First, respondents valued the opportunity to practice language–in a low- stakes, one- on-one context without feelings of judgment. The immediacy and accuracy of AI feedback led to increased confidence and risk-taking, paralleling the quantitatively assessed motivational impacts. Second, students displayed rapid trust and verify behaviors, recognizing that while the AI answers were in general helpful, they had to be double-checked to avoid errors or misinterpretations. This indicates an enhancement of reflective learning approaches and transformative engagement. Third, teachers believed that their pedagogical work had been transformed from instruction to design and facilitation. Digital Facilitators Teachers identified the evolving teacher-student-AI triangle (curation of questions, helping students to judge AI output and promoting strategic interactions as opposed to just simply delivering answers) in today's language classrooms.

Taken together, these results suggest that LLM through language learning input can lead more to repetitive self-motivated practice, increased motivation and decreased anxiety as well as gains in writing accuracy. The findings indicate that AI tools can be effective scaffolds for text production and metalinguistic awareness, even without significant gains in oral fluency. The findings also underscore the importance of teacher help, critical testing and timely quality checking in realizing the educational potential of AI-based language models.

VII. CONCLUSION

The integration of LLMs in language learning settings is empirically justifiable from this research, especially for the enhancement of writing correctness, manuscript fluency and learner motivation. The results illustrate the potential and limitations of LLMs in educational contexts, and they are in line with and add to the literature.

A. Summary of Findings

Treatment students who completed conversational tutoring with LLM-based tutors had significantly greater improvements in writing accuracy, with a mean difference of 6.2 fewer errors per 100 words after adjustment ($p< .01$) based on the quasi-experiment design. Those gains were most noticeable for intermediate (B1) learners, suggesting that LLMs can effectively serve as the scaffold for language output at this level. There were also significant gains on written fluency ($d = 0.29$), but no significant effects for oral fluency, in line with the notion that interactional modality might influence the extent to which learners' fluency is enhanced. The therapy condition likewise received affective benefits with reduced state anxiety and increased intrinsic motivation after the session. These findings underscore the role of LLMs in providing personalized, low-stakes interlanguage practice to enhance engagement and reduce language anxiety.

B. Pedagogical Implications

The implications of the study for language teaching are manifold. First, for the prompt feedback and opportunities for repeated revision that LLMs can offer to learners, they might be valuable tools in personalized practice. This is reminiscent of the role of interaction and negotiation of meaning in language learning, as stressed by interactionist theories of second language acquisition. Second, by allowing teachers to focus on more nuanced pedagogical activities such as content curation and mediating critical interactions with AI feedback, LLMs can complement human instruction. This is reflected in the shift of teacher roles from being traditional lecturers to being facilitators and designers. The study also highlights the need to be cautious in incorporating LLMs to language teaching programmes. In order to prevent students from being overly reliant on the technology and develop their independent thinking, teachers need to equip themselves with methods of training that will enable them use AI technologies adequately.

C. Limitations and Future Research

There are limitations to the study though it is informative. Limitations The usage of one LLM platform might affect the generalisability of the results. Further research should compare the effectiveness of LLMs and explore their impact on different types of learners and learning environments. In addition, further research is needed for long-term effects of LLM-assisted language learning because the current study focused only on short-run results. Longitudinal studies may offer further insight into the sustained impact of LLMs on language quality and learner independence.

D. Conclusion

In conclusion, this study contributes to the growing body of research on using LLMs in language education. It shows how LLMs may invlove teachers in language classroom and facilitate writing accuracy, fluency and learner motivation. By addressing the constraints identified and pursuing further research, teachers may maximise LLMs as effective and engaging agents in language learning.

VIII. REFERENCES

[1] Li, B. (2025). A systematic review of empirical generative AI research in language learning and teaching from 2023 to 2024. *ScienceDirect*. ScienceDirect

[2] Liu, J., & Wang, Y. (2025). How do generative artificial intelligence (AI) tools and large language models (LLMs) influence critical thinking in English as a Foreign Language (EFL) education? *SpringerLink*. SpringerOpen

[3] Meyer, J. (2024). Using LLMs to bring evidence-based feedback into the classroom. *ScienceDirect*. ScienceDirect

[4] Murtaza, M. (2025). The impact of LLM chatbots on learning outcomes in advanced driver assistance systems education. *Nature*. Nature

[5] Sharma, S. (2025). The role of large language models in personalized learning. *SpringerLink*. SpringerLink

[6] Wang, L. (2025). Accuracy of large language models when answering clinical research questions: A network meta-analysis. *JMIR*. JMIR Publications

[7] Luo, X. (2024). Potential roles of large language models in the creation of systematic reviews and meta-analyses. *PMC*. PMC

[8] Chiu, T. (2025). New meta-analysis: LLMs boost cognitive and emotional engagement in groups. *LinkedIn*. LinkedIn

[9] Mizumoto, A. (2025). Large language models fall short in classifying learners' proficiency levels. *ScienceDirect*. ScienceDirect

[10] Idan, D. (2025). Primer on large language models: An educational overview. *PMC*. PMC

[11] Liu, J., & Wang, Y. (2025). The dual nature of generative AI tools and LLMs on critical thinking in EFL education. *SpringerLink*. SpringerOpen

[12] Stadler, H., et al. (2024). Risks of large language models in education. *arXiv*. arXiv

[13] Li, R. (2025). Delving into the practical applications and pitfalls of large language models in education. *PMC*. PMC

[14] Wang, L. (2025). Accuracy of large language models when answering clinical research questions. *JMIR*. JMIR Publications

[15] Murtaza, M. (2025). The impact of LLM chatbots on learning outcomes in advanced driver assistance systems education. *Nature*. Nature

[16] Sharma, S. (2025). The role of large language models in personalized learning. *SpringerLink*. SpringerLink

[17] Meyer, J. (2024). Using LLMs to bring evidence-based feedback into the classroom. *ScienceDirect*. ScienceDirect

[18] Liu, J., & Wang, Y. (2025). How do generative artificial intelligence (AI) tools and large language models (LLMs) influence critical thinking in English as a Foreign Language (EFL) education? *SpringerLink*. SpringerOpen

[19] Li, B. (2025). A systematic review of empirical generative AI research in language learning and teaching from 2023 to 2024. *ScienceDirect*. ScienceDirect

[20] Chiu, T. (2025). New meta-analysis: LLMs boost cognitive and emotional engagement in groups. *LinkedIn*. LinkedIn