COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

AI-Powered Language Learning: Exploring Learner Perceptions and Engagement with Large Language Models in Second Language Acquisition

Diploma Thesis

2025

Oleksandr Lytvyn

COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

AI-Powered Language Learning: Exploring Learner Perceptions and Engagement with Large Language Models in Second Language Acquisition

Diploma Thesis

Study program: Cognitive Science

Field of study: Computer Science

Supervising department: Department of Applied Informatics

Supervisor: PhDr. Ing. Tomáš Gál, PhD.

2025 Oleksandr Lytvyn

60745586

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Oleksandr Lytvyn
Študijný program: kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: AI-Powered Language Learning: Exploring Learner Perceptions and Engagement with Large Language Models in Second Language Acquisition
Výučba jazykov podporovaná umelou inteligenciou: Skúmanie vnímania a angažovanosti učiacich sa pri využívaní veľkých jazykových modelov v osvojovaní si druhého jazyka

Anotácia: Veľké jazykové modely (LLM), ako napríklad ChatGPT, prinášajú v osvojovaní si ďalšieho jazyka, najmä prostredníctvom svojho interaktívneho a nízko-stresového učebného prostredia.

Cieľ: 1. Preskúmať literatúru z oblasti osvojovania si cudzieho jazyka, z kognitívnej vedy a prijímania technológií, so zameraním na afektívne a interakčné rámce.
2. Navrhnúť a realizovať kvantitatívny dotazníkový výskum s aktívnymi užívateľmi LLM pri výučbe jazyka, ktorý meria vnímanú užitočnosť, dôveru, angažovanosť, emocionálne prežívanie a sebahodnotené výsledky učenia.
3. Analyzovať dáta pomocou štatistických metód s cieľom identifikovať kľúčové prediktory angažovanosti a dôvery.
4. Interpretovať zistenia voči pedagogickým teóriám a dizajnu AI nástrojov.

Literatúra: Xiao, Y., & Zhi, D. (2023). ChatGPT in EFL Classrooms: Opportunities and Limitations. ELT Journal.
Camilleri, D. (2024). Trust and Perception in AI-Assisted Language Learning: A Learner-Centered Perspective. Journal of Educational Technology.

Vedúci: PhDr. Ing. Tomáš Gál, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. RNDr. Tatiana Jajcayová, PhD.

Spôsob sprístupnenia elektronickej verzie práce:
bez obmedzenia

Dátum zadania: 30.01.2025

Dátum schválenia: 01.02.2025

prof. Ing. Igor Farkaš, Dr.
garant študijného programu

....
študent

....
vedúci práce

60745586

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Oleksandr Lytvyn
Study programme:	Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study:	Computer Science
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: AI-Powered Language Learning: Exploring Learner Perceptions and Engagement with Large Language Models in Second Language Acquisition

Annotation: Large Language Models (LLMs) such as ChatGPT have introduced new possibilities in second language acquisition (SLA), particularly through low-stress, interactive learning environments.

Aim:
1. Review literature in SLA, cognitive science, and technology acceptance, with a focus on affective and interaction-based frameworks.
2. Design and conduct a quantitative survey targeting active users of LLMs for SLA, measuring perceived usefulness, trust, engagement, emotional affect, and learning outcomes.
3. Perform statistical data analysis in order to identify key predictors of learner engagement and trust.
4. Interpret the findings in relation to pedagogical theory and AI tool design.

Literature:
Krashen, S. D. (1982). Principles and Practice in Second Language Acquisition. Pergamon.
Xiao, Y., & Zhi, D. (2023). ChatGPT in EFL Classrooms: Opportunities and Limitations. ELT Journal.
Camilleri, D. (2024). Trust and Perception in AI-Assisted Language Learning: A Learner-Centered Perspective. Journal of Educational Technology.

Supervisor:	PhDr. Ing. Tomáš Gál, PhD.
Department:	FMFI.KAI - Department of Applied Informatics
Head of department:	doc. RNDr. Tatiana Jajcayová, PhD.
Assigned:	30.01.2025
Approved:	01.02.2025

prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

....

Student

....

Supervisor

Declaration

I hereby declare that the work presented in this thesis is original and the result of my own investigations. Formulations and ideas taken from other sources are cited as such.

Acknowledgment

I would like to begin by expressing my sincere gratitude to my supervisor, PhDr. Ing. Tomáš Gál, PhD., for his continuous support and insightful guidance throughout this project. Working with him was a seamless experience, and I truly appreciated his readiness to help whenever I needed it. This research would not have been possible without the valuable input of Univ.-Prof. Dr. Susanne Maria Reiterer and Mgr. Yuta Watanabe from the Linguistics Department at the University of Vienna. Their advice and encouragement were instrumental in shaping this thesis. Lastly, I'm deeply thankful to my wife for her support, strength, and, most importantly, for lifting my spirits when I needed it most.

Table of Contents

Abstract

Veľké jazykové modely (LLM), ako napríklad ChatGPT, si rýchlo nachádzajú uplatnenie pri osvojovaní si cudzieho jazyka. Dôvody, pre ktoré ich dospelí učiaci sa využívajú dlhodobo, však zostávajú málo preskúmané. V tejto diplomovej práci vychádzame z modelu akceptácie technológií (TAM), ktorý rozširujeme o emocionálne a percepčné dimenzie („TAM-Plus"), a analyzujeme kognitívne a emocionálne faktory dlhodobého využívania LLM. Výskum sme realizovali prostredníctvom dotazníkového prieskumu medzi 101 respondentmi, ktorí používajú veľké jazykové modely na učenie sa cudzieho jazyka.Zistenia potvrdzujú, že respondenti považujú využívanie LLM za užitočné a jednoduché, pričom ich dlhodobá angažovanosť sa prejavuje najmä vtedy, keď prežívajú pozitívne emócie a vnímajú LLM ako dôveryhodného tútora, nie len ako bežný nástroj.
Naše výsledky prispievajú k spresneniu teórie akceptácie technológií a identifikujú vzorec tzv. „kritickej dôvery". Na základe toho navrhujeme, aby boli rozhrania LLM koncipované spôsobom, ktorý podporuje interakcie podobné práci s tútorom a zároveň posilňuje autonómiu používateľa. V závere uvádzame odporúčania pre budúci výskum a návrhy na dizajn, pričom osobitnú pozornosť venujeme multimodálnym a dlhodobým prístupom.

Abstract

Large language models (LLMs) such as ChatGPT are rapidly finding a place in second-language learning, yet we still know little about purpose adult learners keep using them. Grounded in the Technology Acceptance Model and extended with affective and role-perception dimensions ("TAM-Plus"), this thesis surveys 101 language learners who regularly use LLMs to uncover the cognitive and emotional drivers of sustained engagement . Results confirm that ease of use and perceived usefulness remain essential, but lasting commitment emerges when learners feel positive emotions and frame the LLM as a trusted tutor rather than a simple tool . These insights refine acceptance theory, reveal a pattern of "critical trust," and suggest that LLM interfaces should scaffold supportive, tutor-like interactions while promote autonomy. The work ends with design guidelines and research directions for more multimodal, longitudinal studies.

Chapter 1: Introduction

1.1 Background & Rationale

The integration of artificial intelligence (AI) into education is reshaping pedagogical approaches, with second language acquisition (SLA) emerging as a particularly dynamic space for innovation. Among the most impactful developments are Large Language Models (LLMs), like ChatGPT, which represent a significant leap beyond traditional Computer-Assisted Language Learning (CALL) tools. Earlier CALL systems often relied on rigid, decontextualized tasks, where LLMs, built on transformer-based architectures and vast training datasets, offer flexible, human-like interactions that support more personalized and responsive learning (Belda-Medina & Calvo-Ferrer, 2022; Zhang & Huang, 2024). The potential of LLMs aligns well with several core SLA theories, including the Interaction Hypothesis (Long, 1996), Krashen's Input and Affective Filter Hypotheses (Krashen, 1982), Vygotsky's Zone of Proximal Development (1984), and the Input/Output Hypotheses. These frameworks emphasize the role of meaningful interaction, scaffolding, and accessible input in facilitating language acquisition. In this light, LLMs offer opportunities for negotiated dialogue and real-time feedback, reinforcing Long's (1996) view that conversational interaction makes input more comprehensible. Krashen's (1982) emphasis on emotional states is equally relevant: when learners feel engaged, curious or at ease, their affective filter usually lowers, increasing receptivity to input. On the other hand, frustration or anxiety, often coming from impersonal or erroneous AI outputs can block language acquisition, even when linguistic input is otherwise effective.

Recent studies suggest that LLMs may improve low-anxiety environments that promote skill development and learner confidence. Reported benefits include gains in vocabulary, fluency, and written expression, along with a boost in motivation due to the model's accessibility and non-judgmental tone (Xiao & Zhi, 2023; Pratiwi et al., 2024). These findings point to LLMs not only as language learning tools but also as a language partner, helping with emotional facilitating that shape user engagement and perseverance.

Since ChatGPT's release in late 2022, educational settings have seen a surge in enthusiasm, with many users viewing it as a peer-like or tutor-like presence that encourages risk-taking and learner autonomy (Wei, 2024; Pan et al., 2024). This excitement is connected with growing concerns around factual inaccuracies, ethical dilemmas such as plagiarism and data security, and a perceived emotional disconnect in AI-driven interactions (Dam et al., 2024; Theophilou et al., 2023). These issues have led to institutional pushback, driven by fears that reliance on AI could erode critical thinking and authentic communication practices. This tension between the promise of LLMs and their possible drawbacks reveals a critical gap in current scholarship. While technical functionality and theoretical compatibility have been well explored, the learner's subjective experience with AI remains less understood, particularly in terms of emotional engagement and cognitive trust.

As Camilleri (2024) notes, “few studies have critically evaluated the pros and cons of AI-enabled LLMs in educational contexts or examined factors influencing their use” (p. 12). Much of the existing literature privileges measurable outcomes, like vocabulary gains, over affective or perception-driven insights. Adoption frameworks such as the Technology Acceptance Model (TAM; Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT; Venkatesh et al., 2003) offer valuable predictors of behavioral intent, but they don’t fully capture how trust, enjoyment, discomfort, or hesitation influence learner interaction with AI in SLA contexts.

This study addresses that oversight by centering learners’ affective and cognitive responses to LLM use. Through a cross-sectional survey design, it explores how users interpret the role of these tools: as tutors, peers, or neutral assistants, and how factors like trust, emotional resonance, and perceived usefulness shape their engagement. By grounding these human dimensions, the research aims to support language learning and the development of AI applications that are not only pedagogically sound but also emotionally attuned and ethically responsible.

Large Language Models (LLMs) like ChatGPT bring notable promise to second language acquisition (SLA), offering learners interactive, personalized practice, continuous availability, and immediate feedback. These features make LLMs attractive tools for promoting learner autonomy, lowering anxiety, and boosting engagement. Their integration into educational settings however presents a complex picture. While early adopters point to increased motivation and accessible, low-stakes practice opportunities (Pratiwi et al., 2024; Pan et al., 2024), educators and scholars have raised valid concerns about factual inaccuracies (hallucinations), ethical dilemmas (including data privacy and academic integrity) and the risk of over-dependence that may compromise critical thinking (Dam et al., 2024; Theophilou et al., 2023).

Though such issues provide essential context, this study does not aim to resolve the broader ethical or epistemological questions surrounding LLM use. Instead, it is directed into a more specific and underexplored area: learners' perceptions, emotional engagement, and trust development in the context of LLM-supported SLA. Prior research often aligns LLMs with established theories like the Interaction Hypothesis (Long, 1996) and Krashen’s (1982) comprehensible input, yet much of this work focuses primarily on linguistic gains, such as vocabulary development. Similarly, models like the Technology Acceptance Model (TAM; Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT; Venkatesh et al., 2003) help explain user uptake but fall short in capturing emotional nuances such as enjoyment, frustration, or erosion of trust that can significantly shape sustained engagement.

This absence of learner-centered need poses challenges for the thoughtful design and implementation of AI tools in education. Without a clearer understanding of how learners interpret the role of LLMs, whether as tutors, collaborators, or neutral tools educators may adopt technologies that fail to resonate with actual learner needs. Developers, as well, risk focusing on surface-level efficiency while overlooking the deeper features that cultivate trust, reflection, and self-regulation. By investigating learners’ emotional responses, conceptual framing, and trust dynamics, this study seeks to fill that gap—offering a more human-centered and theoretically anchored perspective on how intelligent systems can genuinely support, rather than simply deliver, second language learning.

1.2 Significance of the Study

This study holds significance both in its theoretical insights and its practical implications for second language acquisition (SLA) and educational technology. On the theoretical front, it connects the evolving capabilities of Large Language Models (LLMs) with foundational SLA constructs, offering a quite fresh lens on how AI-mediated interactions intersect with established learning theories. By examining how learners perceive LLMs it explores how real-world use aligns with frameworks like the Interaction Hypothesis, the Zone of Proximal Development (ZPD), and the Input/Output Hypotheses. These connections shed light on how LLMs influence learner engagement, supporting negotiated interaction, self-directed learning, and the reduction of affective barriers such as anxiety or fear of negative evaluation. In doing so, this study contributes to an emerging theoretical model that brings together cognitive, emotional, and behavioral dimensions of AI-supported SLA. From a practical standpoint, the research offers timely, learner-centered perspectives that can guide educators, instructional designers, and ed-tech developers in the responsible integration of LLMs into language learning environments. By foregrounding users' lived experiences, trust dynamics, and emotional responses, it uncovers patterns that can inform more effective pedagogical strategies. These include creating conditions that enhance learner autonomy, designing tasks to improve speaking fluency, engage active use of language, prevent over-reliance, and addressing ethical challenges like misinformation. The findings are also relevant for shaping institutional policies and teacher training initiatives, helping ensure that AI tools are not only technologically robust but also pedagogically meaningful and ethically grounded.

More broadly, this work contributes to ongoing conversations around digital literacy and the future of human-AI collaboration in education. It highlights the nuanced ways learners build trust, maintain engagement, and practice critical reflection when interacting with LLMs—positioning them not as passive recipients, but as active participants in adaptive learning ecosystems. Ultimately, the goal is to support the design of AI-enhanced language learning experiences that are inclusive, transparent, and responsive to the cognitive and emotional needs of diverse learners.

1.3 Interdisciplinary Foundations

This study sits at the intersection of cognitive science, didactics, second language acquisition (SLA), and artificial intelligence (AI), adopting an interdisciplinarity that's essential for unpacking the nuanced interactions between learners and Large Language Models (LLMs). As LLMs become increasingly integrated into educational settings, there's a growing need for frameworks that address their cognitive, pedagogical, and socio-technical dimensions in tandem.

From a cognitive science standpoint, the research draws conceptually on theories of language processing, human-computer interaction (HCI), and socio-cognitive behavior. Through simulated dialogue, LLMs create opportunities for learners to engage in core language learning processes, like real-time negotiation of meaning, explanation of grammar, and self-correction. Observing how learners interact with AI-generated output offers valuable insights into these cognitive behaviors. While constructs such as attention, working memory, and metacognitive regulation inform the theoretical backdrop, it's important to note that these mechanisms are not empirically measured in the current study. Rather, they provide a lens for interpreting learner behaviors, ensuring that the scope and empirical aims of the research are accurately conveyed.

Within applied linguistics and SLA, the study explores how learners perceive and use LLMs through the lens of established pedagogical frameworks. Concepts such as scaffolding, interactional competence, and self-regulated learning help contextualize both the learners' behavior and their reflections on AI feedback. This dimension grounds the use of advanced technology in well-established understandings of language learning and educational practice. The study also engages with perspectives from computer science and AI ethics, critically examining how LLMs are designed and deployed within educational environments. It considers key issues such as algorithmic transparency, trust calibration between humans and machines, and the preservation of learner agency particularly when instructional functions like feedback are delegated to AI. These discussions contribute to broader ethical debates on the responsible integration of intelligent systems in classrooms.

By putting together theories and methods from these complementary disciplines, this research offers a pedagogically sound, and ethically reflective examination of LLM use in SLA. It aims to deepen our understanding of how human learners interact with, adapt to, and potentially co-evolve with intelligent technologies in digitally mediated learning contexts.

1.4 Thesis Structure

To provide a clear roadmap for the reader, this thesis is organized into six chapters, each building sequentially to explore learner engagement with Large Language Models (LLMs) in the context of second language acquisition (SLA):

Chapter 1: Introduction

Current chapter lays the foundation by presenting the study's background, research problem, significance, interdisciplinary framing, and objectives, emphasizing the importance of understanding learner perceptions of LLMs from both theoretical and practical perspectives.

Chapter 2: Literature Review

Explores the evolution of AI in language education, with a particular focus on the role of LLMs in SLA. It integrates classical SLA theories with digital pedagogical approaches and critically examines existing research on learner attitudes, benefits such as motivation and fluency, and challenges like trust and hallucinations. This chapter gives a theoretical background for the research, helping address existing gaps.

Chapter 3: Methodology

Details the cross-sectional, quantitative research design, including participant selection, data collection instruments rooted in TAM and affective models, and procedures for recruitment, ethical safeguards, and statistical analysis.

Chapter 4: Results

Presents the survey findings through descriptive statistics, correlation analyses, and regression outputs. It highlights learner's patterns of LLM use, trust indicators, and perceived cognitive and emotional outcomes.

Chapter 5: Discussion

Interprets the results in light of SLA theory and cognitive science, drawing comparisons with existing studies. It considers implications for pedagogy, AI tool design, and learner autonomy within digitally mediated learning environments.

Chapter 6: Conclusion

Summarizes the key findings and theoretical contributions, reflects on practical applications, and offers directions for future research into LLM use in SLA, particularly from learner-centered and interdisciplinary perspectives.

Chapter II: Literature Review

2.1 AI and Large Language Models in Second Language Acquisition (SLA)

2.1.1 From CALL to LLM-Enhanced Tools

The integration of AI into second language acquisition marks a shift from static, rule-based Computer-Assisted Language Learning (CALL) systems to more dynamic tools, including task-specific chatbots and, more recently, transformer-based Large Language Models (LLMs) like ChatGPT or DeepSeek. Early CALL applications were often limited by their lack of authentic dialogue (Belda-Medina & Calvo-Ferrer, 2022). Chatbots offered incremental improvements, however their reliance on pre-programmed responses restricted adaptability until advances in deep learning and natural language understanding made more flexible interactions possible (Belda-Medina & Calvo-Ferrer, 2022). Today's LLMs enable real-time, personalized dialogues, support multimodal communication, and are grounded in vast training datasets (Zhang & Huang, 2024; Park, 2023). This evolution reflects a broader pedagogical turn from rigid, vocabulary drills to immersive environments that promote learner agency, critical thinking, and autonomy (Xiao & Zhi, 2023).

2.1.2 Key Findings from Recent Studies

Recent empirical studies highlight LLMs' potential in vocabulary acquisition. In Zhang and Huang's (2024) controlled experiment, learners using LLM chatbots showed vocabulary gains of 7.6% in productive and 5.0% in receptive measures. A delayed post-test suggested an even greater advantage—12.39%—though questions remain about whether the study controlled for external exposure or tracked long-term retention beyond the 12-week period, both of which are essential for assessing lasting impact. Notably, 69% of participants acquired incidental, non-target vocabulary during interaction, pointing to the organic learning potential of LLM-based systems. However, these results warrant cautious interpretation, as the study does not disclose variability in sample size or control for external learning influences. Ngo (2024) links learners' engagement to adaptive feedback mechanisms aligned with the i+1 principle and spaced repetition. Still, most reported metrics such as a 76.92% usage rate—rely on self-reports rather than objective behavioral data or performance-based assessment. Future research would benefit from outcome-based measures to more accurately evaluate instructional value.

2.1.3 Beyond Vocabulary: Writing, Speaking, and Engagement

Beyond lexical learning, LLMs also influence writing and speaking skills. For writing, ChatGPT supports brainstorming and revision through genre-aware suggestions and language framework (Xiao & Zhi, 2023), though learners often need to critically assess the tool's occasional inaccuracies. In speaking practice, LLM-driven interaction has been linked to reduced anxiety and measurable improvements in fluency (B = 0.65), vocabulary (B = 0.57), and accuracy (B = 0.46), thanks to adaptive feedback and adjustable difficulty levels (Qiao & Zhao, 2023). While these figures are promising, they invite scrutiny regarding sample size and the sustainability of learning gains post-intervention.

As for learner engagement, LLMs promote motivation and autonomy by tailoring dialogue topics to individual interests (Ji et al., 2022; Park, 2023). However, much of this research is grounded in learner perceptions, with limited mapping through behavioral or longitudinal data. The literature often skews toward optimism, at times conflating enjoyment with actual pedagogical effectiveness.

2.2 SLA Theories & AI

Large Language Models (LLMs) are increasingly interpreted through foundational SLA theories that emphasize interaction, scaffolding, input, and output. This section evaluates how four core frameworks: Long's Interaction Hypothesis, Vygotsky's Zone of Proximal Development (ZPD), Krashen's Input Hypothesis, and Swain's Output Hypothesis are reflected in LLM-mediated learning, while also addressing key theoretical limitations.

Interaction Hypothesis: long's Interaction Hypothesis posits that language learning is driven by meaningful exchanges in which learners negotiate meaning (Long, 1996). LLMs simulate this through turn-based, responsive dialogue. Ji et al. (2023) highlight that conversational AI can establish "an authentic environment for communication in a target language" (p. 2), with tools like ChatGPT prompting clarification and reformulation—thus encouraging interaction aligned with this framework (Ngo, 2024). However, the quality of such exchanges can be compromised by NLP or ASR errors, which interrupt communicative flow (Ji et al., 2023, p. 9). More critically, LLMs lack the reciprocity inherent in real social interactions, limiting their ability to replicate the dynamics of peer conversation.

Zone of Proximal Development (ZPD): According to ZPD theory, learning occurs when guidance from a more knowledgeable other helps a learner progress from what they can do independently to what they can achieve with support. ChatGPT can mimic scaffolding by adjusting prompts and feedback to a learner's level (Qiao & Zhao, 2023). Ji et al. (2023) describe how AI systems classify users into proficiency tiers to deliver adaptive support. Still, comparing this functionality with human scaffolding overlooks key elements, particularly emotional and interpersonal nuances. As Xiao and Zhi (2023) note, the "absence of genuine social reciprocity" continues to be a significant limitation.

Input and Output Hypotheses: Krashen's Input Hypothesis centers on the value of comprehensible input that stretches the learner just beyond their current competence (i+1). LLMs appear to fulfill this role by embedding new vocabulary and structures within contextually appropriate responses (Ngo, 2024). Complementing this, Swain's Output Hypothesis stresses the importance of language production for internalization. LLMs offer low-pressure opportunities for learners to generate language, reinforced by nonjudgmental feedback and adaptable correction mechanisms (Ji et al., 2023, p. 10; Zhang & Huang, 2024). However, much of the existing evidence is based on self-reporting rather than performance metrics, signaling the need for theory-driven, pre-registered research that rigorously tests these claims.

Synthesis and Theoretical Critique: although LLM capabilities are often mapped onto SLA theories, this alignment is frequently retrospective. Many studies adopt a post hoc approach by matching LLM features to existing models without prospective hypotheses or experimental validation. While theoretical resonance with Interaction, ZPD, Input, and Output frameworks is evident, pedagogical equivalence remains largely speculative. Moving forward, hypothesis-based longitudinal research is essential to examine these relationships more robustly, particularly with respect to affective, interpersonal, and contextual dimensions that are difficult for AI to replicate.

2.3 Learner Attitudes Toward AI

Learners' perceptions of AI-based language tools are shaped by a mix of trust, usefulness, and emotional responses. Frameworks like the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Technology Acceptance Model (TAM) conceptualize adoption in terms of expected performance, perceived effort, and social influence. Camilleri (2024) highlights that trust, especially in the reliability of AI sources is a strong predictor of performance expectancy ($\beta = 0.450$, $p < 0.001$), suggesting that credibility plays a key role in learners' willingness to engage with these tools.

Positive Attitudes: Excitement and Engagement: many students describe their experiences with AI as interactive and supportive, often highlighting the personalized feedback and multimodal features of platforms like Replika and ChatGPT (Belda-Medina & Calvo-Ferrer, 2022; Xiao & Zhi, 2023; Wei, 2024). In Qiao and Zhao's (2023) study, learners reported finding AI instruction more enjoyable than peer-led tasks, linking it to gains in speaking ability. These perceptions echo UTAUT's emphasis on performance expectancy, though findings tend to rely on self-reported impressions rather than measured outcomes.

Negative Attitudes: Anxiety and Skepticism: despite these advantages, reservations remain common. Learners voice concerns about AI-generated inaccuracies, its inability to interpret nuanced context, and the emotionally sterile nature of machine interaction (Wei, 2024; Belda-Medina & Calvo-Ferrer, 2022). Data privacy worries also emerge, especially when platforms request access to personal information. While some studies report reduced anxiety in AI-facilitated tasks (Qiao & Zhao, 2023), others document lingering fears about fluency and pronunciation, indicating that performance-related stress isn't fully alleviated by AI support.

Mixed Emotions and Ambivalence: attitudes toward AI often blend enthusiasm with hesitation. Some learners appreciate the confidence boost AI provides (Wei, 2024), yet still question its limitations particularly the absence of human nuance crucial for deep language learning (Xiao & Zhi, 2023). Even when users benefit from AI features, many continue to favor human interaction, which tempers widespread adoption (Belda-Medina & Calvo-Ferrer, 2022).

Measurement Limitations: most studies on learner attitudes use Likert-scale surveys, which offer limited insight into how students actually interact with AI tools. This highlights the need for methodological triangulation by incorporating behavioral analytics like usage logs, time-on-task metrics, and performance outcomes. Additionally, the "novelty effect" such as an spike in early enthusiasm that may decline over time, it remains largely unexplored. Longitudinal research is needed to determine whether positive attitudes persist or fade.

2.4 Reported Benefits

The use of Large Language Models (LLMs) in second language acquisition (SLA) has revealed a range of pedagogical benefits, notably in areas such as personalization, real-time feedback, and learner engagement. These strengths closely reflect core SLA principles, especially those concerning scaffolding, interactivity, and learner autonomy.

Personalization: LLMs tailor content delivery to match learners' proficiency levels, goals, and learning pace (Pan et al., 2024; Xiao & Zhi, 2023). In Xiao and Zhi's study, one participant described ChatGPT as a "peer tutor" that guided them through structured writing activities using customized prompts (p. 5). This adaptability allows learners to shift between directive and collaborative modes, with some tools even recommending materials aligned with curriculum standards (Kim & Kim, 2020). Such dynamic responsiveness mirrors the principles of the Zone of Proximal Development (ZPD) and supports cognitive load management (Ngo, 2024). Nonetheless, questions remain about the long-term viability and equity of scaled personalization, especially when it relies on limited research credits or premium access.

Feedback: one of the most prominent advantages is the immediacy and adaptability of feedback. Ji et al. (2023) describe systems that fine-tune input based on semantic and syntactic accuracy, offering emotionally secure spaces for practice (p. 11). Similarly, Duolingo's AI provides instant corrections on grammar and fluency (Qiao & Zhao, 2023), and Moore and Tsay (2024) report better revision outcomes in LLM-assisted writing tasks. These findings support Swain's Output Hypothesis, as feedback encourages learners to notice and address interlanguage gaps. However, many benchmark studies assess feedback performance using brief prompts of fewer than 40 tokens—raising concerns about their applicability to more complex writing or sustained discourse.

Engagement: LLMs enhance learner engagement through role-play scenarios, personalized interaction styles, and opportunities for exploratory dialogue. Dam et al. (2024) describe chatbots as “objects-to-think-with,” supporting reflection through Socratic questioning (p. 9). Other studies report high motivation levels, as learners practice “anytime and anywhere” (Qiao & Zhao, 2023, p. 10), reinforcing habits of self-regulated learning. Even when interactions are imperfect, the fluidity of conversation tends to foster confidence more effectively than grammatical precision alone, helping to reduce inhibition (Belda-Medina & Calvo-Ferrer, 2022, p. 12).

Theoretical Integration: these reported benefits align closely with established SLA frameworks: personalization reflects ZPD theory, feedback supports Swain’s model, and interaction-fueled engagement resonates with the Interaction Hypothesis. Yet, scaling these benefits in real-world settings will require careful consideration of access, cost, language limitations, and alignment with instructional goals.

2.5 Challenges with LLMs

Despite their potential, large language models (LLMs) bring a host of technical, behavioral, and ethical challenges that complicate their integration into second language acquisition (SLA). Key issues include hallucinated content, growing dependence in academic contexts, and unresolved concerns about bias, transparency, and equitable access.

Technical Limitations and Hallucinations: one of the most pressing concerns is the generation of hallucinations responses that sound plausible but are factually incorrect. These inaccuracies can misinform learners about language rules, cultural norms, or vocabulary, especially when AI-generated feedback is accepted without question (Theophilou et al., 2023). While developers claim improvements in accuracy through Reinforcement Learning from Human Feedback (RLHF), these claims are difficult to verify due to the closed nature of training data and proprietary model architectures (Dam et al., 2024). Additionally, many LLMs rely on outdated data and lack robust logical reasoning, which limits their reliability for nuanced or context-sensitive instruction.

Over-Reliance and Academic Integrity: the widespread availability of LLMs has raised concerns about over-reliance, particularly in academic settings. Increasingly, students turn to generative AI for writing assignments or test responses, often without proper attribution posing risks to academic integrity (Dam et al., 2024). Beyond issues of plagiarism, such use may hinder the development of essential skills like critical thinking and independent reasoning (Theophilou et al., 2023). In response, some institutions have imposed blanket bans on LLM use. However, these restrictions can overlook the legitimate accessibility functions AI provides, including support for students with dyslexia or speech-to-text needs. A more balanced strategy is needed one that safeguards academic standards without excluding learners who rely on inclusive technologies.

Ethical and Emotional Concerns: LLMs also raise unresolved ethical and affective concerns. Their outputs are shaped by opaque training datasets, often embedded with unknown biases that may marginalize non-dominant dialects or cultural identities (Belda-Medina & Calvo-Ferrer, 2022). Described as “black boxes” by Dam et al. (2024), these systems offer little transparency regarding their decision-making processes, making it difficult to ensure fairness or audit outcomes. On the emotional side, while learners often seek relational interaction, current LLMs struggle to interpret tone or convey empathy (Pan et al., 2024). Without the ability to respond to emotional cues or replicate human connection, these tools may fall short of promoting sustained engagement or may even misrepresent cultural sensitivities.

2.6 Digital Literacy and Prompt Training

As large language models (LLMs) become more embedded in language education, the ability to craft effective prompts, known as a “prompt literacy” has emerged as a vital digital skill. Maloy and Gattupalli define it as the capacity to “communicate with and direct generative AI systems without needing expertise in computer programming” (n.d., p. 211). This accessibility helps democratize AI-supported learning, enabling a wide range of students to engage meaningfully with interactive language tasks.

Developing prompt literacy supports more intentional and ethical use of AI tools. By learning to frame precise, goal-oriented inputs and critically assess responses, learners are better equipped to avoid issues like hallucinated content or irrelevant elaboration (Maloy & Gattupalli, n.d., p. 212). Structured training plays a key role here—but staying current is a challenge. Frequent model updates often disrupt established strategies, making it difficult for curricula to keep pace. Without agile instructional design, there's a risk that lessons quickly become outdated.

Although prompt literacy is increasingly framed as a transferable skill, its broader academic impact, particularly on non-AI writing tasks like argumentation, coherence, or genre control remains underexplored. Its cognitive implications within SLA contexts, especially regarding long-term language development and metacognitive awareness, have yet to be thoroughly studied.

This skill also intersects with broader models of digital fluency. Eshet-Alkalai's five-part framework includes reproduction literacy (creative reuse of content) and branching literacy (nonlinear, adaptive thinking)—both of which parallel the iterative nature of prompt engineering (Eshet-Alkalai, 2004, 2005; Aviram & Eshet-Alkalai, 2006).

In sum, prompt literacy is increasingly central to navigating LLM-based learning environments. However, its educational value depends on flexible curricula and further research into how it transfers across domains and supports sustained language development.

2.7 Instructor Mediation and Teacher Roles in AI-Supported SLA

The integration of AI tools like ChatGPT into second language acquisition (SLA) has redefined, rather than replaced, the role of educators. Instead of rendering teachers obsolete, AI shifts their responsibilities toward guiding, contextualizing, and enhancing its educational potential within sound pedagogical frameworks.

Increasingly, instructors are adopting ChatGPT as a teaching assistant for lesson planning, content creation, and formative assessment and even assessments validation. According to Li et al. (2024), this streamlining of routine tasks allows teachers to dedicate more time to individualized feedback and cognitive scaffolding (pp. 10–11). When trained effectively, educators can manage AI use in ways that preserve human judgment and foster emotional connection in the classroom.

This coordination, however, requires specific preparation. As Li et al. (2024) point out, "teachers need specialized training to effectively use AI tools" (p. 12). Without technical competence and pedagogical confidence, aligning AI integration with communicative teaching goals becomes a challenge. Echoing this, Songsiengchai (2025) emphasizes the importance of formal training programs that help teachers embed AI responsibly in instructional contexts (p. 304).

Emerging evidence indicates that educator-led AI use tends to foster deeper learner engagement. In Songsiengchai's (2025) study, students co-designed prompts for ChatGPT to generate practice dialogues (p. 307), actively participating in the learning process. Subsequent focus group interviews revealed increased motivation, engagement, and appreciation for personalized feedback (pp. 308–310). However, much of the current data stems from teacher surveys and anecdotal classroom accounts. To move the field forward, quasi-experimental research is needed to directly compare outcomes between teacher-mediated and AI-exclusive learning environments.

AI support extends across language domains as well. Teachers observed improvements in vocabulary, grammar, and reading comprehension (Songsiengchai, 2025, p. 307), while students found the experience "highly engaging" due to ChatGPT's responsive and tailored interactions (p. 310). Yet these benefits hinge on instructor mediation—on how well teachers guide learners, interpret AI outputs, and promote reflective learning. Without this scaffolding, generative tools may confuse more than they clarify.

In short, teachers remain central to the educational impact of LLMs. Their role in mediating, contextualizing, and humanizing AI use is what ultimately shapes its ethical application and instructional success. Future research must move beyond perception-based findings to examine how educator involvement directly influences learning efficacy in AI-enhanced SLA settings.

2.8 Longitudinal and Developmental Perspectives on LLM Use

Longitudinal Studies: most current research on Large Language Models (LLMs) in second language acquisition (SLA) due to the novelty of technology centres on short-term interventions or learner perceptions at a single point in time. However, recent studies have begun to explore the longer-term impact of sustained engagement with LLMs offering a clearer picture of how language proficiency, motivation, and autonomy evolve over time.

A notable example is the extended intervention conducted by Songsiengchai (2025), where "the experimental group interacted with Chat GPT for 30 minutes daily over eight weeks, while the control group continued with their regular English curriculum" (p. 307). This setup allowed for comparisons not only in immediate outcomes but also in developmental patterns. Results showed strong gains in language performance, supported by a "paired sample t-test reveal that this difference was statistically significant ($p < .001$)" (Songsiengchai, 2025, p. 303). Complementing these quantitative findings, qualitative data provided deeper context. Focus group interviews revealed that "students in the experimental group found the AI-based learning experience more engaging and personalized" (Songsiengchai, 2025, p. 309), while classroom observations recorded signs of emotional growth, including "high motivation, strong confidence, and a positive attitude change" (p. 309). These patterns suggest that consistent exposure to ChatGPT may foster not just linguistic improvement but also meaningful shifts in learner disposition and self-efficacy. These outcomes carry implications beyond individual classrooms. As Songsiengchai (2025) observes, "these findings urge educational policymakers and institutions to explore the promising role of AI in the future of language teaching... A longitudinal study could track the progress and retention of language proficiency among students who have undergone extended AI-assisted language learning programs" (p. 311). Long-term research like this is crucial to understanding the durability of AI-mediated learning and the pedagogical conditions that best support it.

Cross-Cultural Comparisons: beyond the temporal dimension, the success of LLMs in language education is also influenced by cultural and linguistic context. While LLMs are designed for general applicability, their educational impact depends on how well they align with local pedagogical traditions, curriculum frameworks, and cultural attitudes toward technology. As mentioned previously, the models might be biased and the nature of their black box remains present. Songsiengchai (2025) calls for more attention to these variables, suggesting that "comparative studies between Thai students and those from other regions could reveal potential variations in learning outcomes and preferences" (p. 311). Factors such as educational norms, language exposure, and institutional expectations likely shape how learners engage with and evaluate AI tools.

This cultural responsiveness also applies to the tools themselves. As Songsiengchai (2025) explains, "a collaboration between educational institutions, AI developers, and language experts is crucial in designing and developing AI platforms that are culturally sensitive and aligned with the Thai curriculum" (p. 304). Without such alignment, even the most advanced tools may fall short of supporting the goals of local educators and learners.

This is particularly important in contexts where traditional language teaching methods still dominate. As highlighted, "this integration must be addressed with care to ensure that it supports traditional teaching approaches and enhances the language-learning experience" (Songsiengchai, 2025, p. 311). Without careful adaptation, the introduction of LLMs could clash with established norms, reducing their educational value.

Together, these longitudinal and cross-cultural insights provide a fuller understanding of how learners grow with LLM tools over time and how this growth varies by context. By focusing on both development and diversity, research can move beyond static snapshots to support more sustainable, equitable models of AI-enhanced language learning.

Chapter Synthesis: the analysis of longitudinal studies highlights critical insights about sustained learner engagement, skill retention, and developmental trajectories in LLM-mediated SLA. These temporal and cultural insights clarify essential conditions for meaningful and sustainable integration of AI. Bringing together all previous discussions such as theoretical alignment, learner attitudes, digital literacy, instructor mediation, and longitudinal outcomes, the final section synthesizes the identified insights and critical gaps, positioning the current research within a clear, well-defined conceptual framework.

2.9 Summary and Research Gap

Although the role of large language models (LLMs) in second language acquisition (SLA) has been increasingly explored, several key gaps continue to limit both theoretical understanding and practical implementation. This review highlights three pressing areas that warrant further investigation:

Learner agency and prompt literacy: Most existing studies portray learners as passive recipients of AI output, giving little attention to how they actively shape interactions through prompting. There is a clear need to reconceptualize prompt literacy as a language-related skill, one that plays a central role in co-constructing feedback, building trust, and sustaining cognitive engagement.

Instructor mediation mechanisms: While teacher involvement is often emphasized, few studies provide empirical insight into how instructors influence learners' attitudes, emotional reactions, or their evaluation of AI-generated input. Clarifying these mechanisms would support more effective integration of AI tools within communicative language teaching.

Sociocultural equity and contextual responsiveness: Research rarely accounts for how factors such as socioeconomic background, institutional infrastructure, or local pedagogical norms shape the accessibility and impact of LLMs. Future studies must move toward more inclusive, context-sensitive models that reflect real-world diversity.

While other areas such as affective outcomes and developmental trajectories also remain underexamined, this study specifically targets the first gap: how learners assert agency and develop prompt literacy in shaping AI-supported language learning. This focus informs the cross-sectional, quantitative survey design detailed in Chapter 3, which maps learner perceptions, behaviors, and emotional responses across a diverse user base.

Chapter 3: Methodology

3.1. Research Design

This study adopts a quantitative, cross-sectional research design, selected to address core questions about language learners' perceptions, emotional engagement, and interaction patterns with Large Language Models (LLMs) in second language acquisition (SLA). This approach captures a snapshot of language learner attitudes and behaviors at a single point in time, offering timely insights into how LLMs are currently being experienced and conceptualized by users. The quantitative framework was chosen for several reasons. Primarily, it supports statistical analysis and provides measurable results in SLA contexts. Using structured surveys grounded in models such as the Technology Acceptance Model (TAM) ensures methodological precision in measuring constructs like perceived usefulness, trust, ease of use, and behavioral intention. This level of clarity enables the generation of empirically supported recommendations for both educators and developers.

The decision to employ a cross-sectional rather than longitudinal or qualitative method reflects both practical and theoretical considerations. Although longitudinal designs are valuable for tracking shifts over time, they require extended timelines and participant retention, which can be difficult given the pace of technological change in generative AI. Likewise, qualitative or mixed-method approaches, while effective for exploring individual learner narratives and contextual nuance, often involve smaller sample sizes and more resource-intensive data collection limiting their scalability and applicability for large-scale analysis. In contrast, this study addresses those limitations by validating earlier qualitative findings through quantitative means. Prior research has identified affective and perceptual themes (e.g., Camilleri, 2024; Xiao & Zhi, 2023), but this study extends those insights across a broader participant pool, systematically mapping learner trust, the ways they perceive AI, emotional response, and engagement patterns at scale.

It's important to note that while this design allows for the identification of correlations and predictive relationships, it does not establish causality. Regression analyses show statistical associations, but follow-up longitudinal research would be needed to trace how these relationships evolve over time.

Overall, the chosen methodology strikes a balance between empirical accuracy and practical feasibility. It offers a scalable, theory-informed framework for capturing the learner experience with LLMs by directly responding to calls for large-scale, data-driven insights in the evolving field of AI-mediated language learning.

3.2 Participants

This study draws on a convenience sample of approximately 100 to adult second-language learners who regularly use Large Language Models (LLMs), such as ChatGPT, for language learning. To qualify, participants must self-report using an LLM at least once a week for activities like conversational practice, vocabulary building, grammar correction, or pronunciation support. This criterion ensures that respondents have meaningful, hands-on experience and can offer informed perspectives on AI-assisted learning.

Participants will be recruited through a multi-channel outreach strategy aimed at reaching a digitally literate and linguistically diverse population. Recruitment platforms will include: (a) subreddit communities such as r/languagelearning, (b) Facebook groups and similar social media forums focused on language learning, and (c) Poster with QR code for employees from various departments within a private international company.While these channels widen demographic reach, the sample is probably inclined to attract self-selecting individuals, especially those who are tech-savvy, inherently motivated, and receptive to innovative methods of language acquisition. This represents a recognized type of self-selection bias, in which participants are not randomly selected from the wider learner community, but rather choose to join the study due to existing traits or interests that correspond with the research setting. Due to practical limitations, convenience sampling, often from the researcher's own academic or professional networks, remains the standard in psychological and educational technology research, where this kind of bias is particularly common. The sample may therefore overrepresent the opinions and experiences of early adopters, which could inflate the usefulness or allure of AI-mediated language learning resources. Although these individuals offer insightful information on new usage trends and affective engagement, their viewpoints might not apply to learner populations who are less tech-savvy or more averse. The proper interpretation of results requires acknowledging this constraint, which also emphasizes the necessity for future studies with larger, more demographically and attitudinally varied cohorts.

Since the primary goal is to explore perceptions and emotional responses among current LLM users, no control group of non-users will be included. This design choice supports detailed analysis of variables such as perceived usefulness, trust, enjoyment, and usage patterns within a clearly defined user group. However, it limits generalizability, and the results should not be assumed to represent the broader population of language learners particularly those who do not engage with AI-based tools.

The research employs a cross-sectional survey design to capture current attitudes and self-reported outcomes. While this offers a valuable snapshot of how LLMs are presently used in language learning, it does not support causal claims or provide insight into long-term developmental changes. These methodological boundaries are acknowledged, and findings will be interpreted as descriptive associations rather than evidence of cause-and-effect relationships.

3.3 Data Collection Instruments

To explore learner perceptions, engagement patterns, and emotional responses related to the use of large language models (LLMs) in second language acquisition (SLA), this study employed a structured online survey administered through Google Forms, titled *AI-Powered Language Learning: Learner Perceptions Survey*. The instrument was designed to capture a range of cognitive, behavioral, and affective variables, and was grounded in established theoretical frameworks, specifically the Technology Acceptance Model (TAM; Davis, 1986) and the Unified Theory of Acceptance and Use of Technology (UTAUT).

The survey opens with a consent form, which clearly outlines the voluntary and anonymous nature of participation, in compliance with ethical research standards. Contact information for both the researcher and academic supervisor is provided to promote transparency and accountability.

Participants proceed through a primarily Likert-scale questionnaire, where they rate statements on a five-point scale ranging from "Strongly Disagree" to "Strongly Agree." These items are designed to assess subjective perceptions of the utility, ease of use, trustworthiness, and emotional impact of AI-based language tools, following standard psychometric conventions to ensure both reliability and validity.

The questionnaire is composed of ten distinct sections, each targeting a specific analytical purpose. It begins with the consent form and is followed by a screening question to verify that participants have prior experience with tools such as ChatGPT or DeepSeek, ensuring relevance and consistency across responses. The next section gathers demographic data, including age, gender identity, native and target languages, and self-assessed proficiency levels, allowing for subgroup analysis and improving the ecological validity of the findings. Subsequently, participants report their usage patterns with different LLM platforms, such as ChatGPT, DeepSeek, Gemini and Duolingo Max indicating how often and for how long they use these tools, as well as the types of language-learning tasks they engage in, including writing, speaking, grammar, vocabulary, translation, and pronunciation. These details help contextualize participants' attitudes by linking perceptions to practical usage. The following section probes how learners perceive the role of LLMs in their language development. Here, respondents reflect on whether they conceptualize the tools as tutors, peers, or neutral assistants, and rate items related to trust and perceived reliability. This is followed by a section directly aligned with TAM and UTAUT constructs, which includes measures of perceived usefulness, ease of use, behavioral intention, and social influence—core components of the study's theoretical foundation. To assess affective and motivational aspects, the survey adapts items from the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988), targeting variables such as anxiety reduction, confidence, enjoyment, and motivation. These items were pilot-tested with 30 participants, yielding strong internal reliability (Cronbach's $\alpha = .85$), and an exploratory factor analysis confirmed a coherent single-factor structure. Nonetheless, because all measures are self-reported, their validity remains limited. Future studies would benefit from incorporating objective engagement metrics to strengthen these findings.

Participants are also asked to self-assess their learning outcomes in vocabulary, grammar, writing fluency, speaking confidence, and listening comprehension. While these responses provide insight into perceived progress, they are based on a single measurement method, therefore susceptible to biases such as overestimation or positivity effects. As such, these findings are interpreted as correlational rather than definitive proof of learning effectiveness. The research acknowledges this limitation and encourages future research to employ measures such as pre-/post-tests, blind-rated writing samples, or usage analytics (e.g., frequency and accuracy of error correction). At the end of the survey, participants have the option to provide open-ended feedback regarding their experiences with LLMs. These responses are analyzed using principles of template analysis (King, 2004) to surface recurring themes and contextual insights. Although not included in the formal quantitative analysis, this qualitative component adds interpretative depth and is referenced in the Discussion chapter to explore notable trends or outliers. The final page of the survey thanks participants for their time and reiterates the availability of contact details for those interested in receiving study results. This final step reinforces ethical engagement and fosters participant trust. To address potential common method bias, the survey design includes clearly separated thematic sections, neutral language, and reverse-coded items where applicable. Nevertheless, the self-report format inherently limits the robustness of findings, reinforcing the need for future studies to adopt multi-method designs that incorporate behavioral or performance-based measures.

Recent literature supports the methodological approach taken here. Studies by Losi et al. (2024) and Akhiat (2024) have confirmed the reliability of similar Likert-based instruments in assessing learner perceptions of tools like ChatGPT, particularly regarding constructs such as usability, usefulness, and emotional response. These findings lend further credibility to the design choices of the current study.

By combining rigorously structured quantitative measurement with thoughtfully integrated open-ended responses, the instrument enables both the identification of generalizable trends and the exploration of individual learner experiences, offering a nuanced perspective on how LLMs are currently shaping SLA practices.

.

3.4 Research Procedure

The research procedure was carefully structured to ensure methodological accuracy, ethical integrity, and operational efficiency, following established protocols for quantitative SLA research involving digital tools (Akhiat, 2024). The process unfolded in four stages, beginning with ethical preparation and pilot testing. Prior to data collection, approval was obtained from the supervisor at Comenius University, guaranteeing adherence to human subjects research guidelines. The informed consent form was embedded at the start of the survey, requiring participants to confirm their voluntary participation and understanding of anonymized data use. To refine the instrument, a pilot study was conducted with ten language learners who met the inclusion criteria. Their feedback led to minor revisions in Likert-scale wording and corrections to branching logic errors. This iterative refinement process reflects best practices in survey development and aligns with Akhiat's (2024) recommendation to "optimize clarity and functionality in digital survey instruments" (p. 36). Participant recruitment and survey dissemination followed a multi-channel approach designed to reach a linguistically diverse, technologically literate population of active LLM users. Over a three-week period, the survey was distributed through several targeted platforms: subreddit communities such as r/languagelearning, Facebook groups focused on language education, and university mailing lists associated with language departments. Additionally, social media advertisements were placed on platforms like Instagram and Facebook to expand outreach. To complement digital strategies, a poster campaign was implemented within a large international technology company. Posters featuring a QR code and survey link were placed in communal spaces such as cafeterias and bulletin boards. This strategy enabled access to a professional demographic already familiar with AI tools and engaged in informal or self-directed language learning. The recruitment design mirrors that of Monika and Suganthan's (2024) enterprise-based approach, as documented in Akhiat (2024, p. 49), and was intended to maximize both participant relevance and demographic breadth. Data collection was conducted via Google Forms. The survey began with the consent form and was followed by three main sections. The first gathered demographic and usage data, including age, native and target languages, and frequency and duration of LLM use. The second consisted of Likert-scale items measuring constructs drawn from TAM and engagement frameworks namely, perceived usefulness, ease of use, trust, anxiety reduction, and behavioral intention. The final section offered two optional open-ended questions, inviting participants to describe perceived benefits and challenges of LLM use. These qualitative responses were excluded from the statistical analysis but provided supplementary interpretive context. To reduce dropout rates, the survey was designed to be completed in 10 to 12 minutes, and Google Forms' autosave functionality helped accommodate interruptions. Daily monitoring ensured technical functionality, and weekly reminders were posted across recruitment platforms to sustain engagement.

Following the collection period, data handling protocols ensured anonymity and security. No IP addresses or identifying metadata were recorded, and all submissions were stored in a password-protected Google Drive folder accessible only to the researcher. Duplicate entries—identified through timestamp comparison—and responses with less than 70% completion were excluded. The cleaned dataset was exported in CSV format for compatibility with statistical packages in Python, including pandas and scipy.stats, facilitating further analysis. The decision to exclude follow-up interviews or qualitative expansions prioritized the study's scalability and replicability, reflecting a pragmatic research paradigm focused on delivering "actionable insights through streamlined, real-world data collection" (Akhiat, 2024, p. 49).

Throughout the process, ethical considerations were interwoven into the study's design, particularly around trust, data privacy, and learner autonomy. Participants were informed about how their data would be used and protected and were given the freedom to withdraw or skip any question. Additionally, the survey encouraged participants to critically reflect on their use of AI, reinforcing their role as active agents in the learning process. These elements underscored the study's broader commitment to ethical integrity and responsible AI use in language education, ensuring that both procedural and philosophical aspects of the research aligned with current standards for ethical educational technology implementation.

3.5 Data Analysis

The data analysis approach for this study is carefully designed to address the research questions and hypotheses through a structured combination of descriptive, inferential, and advanced statistical methods, including mediation analysis. All analyses were conducted using Python's data science libraries, chosen for their flexibility, transparency, and alignment with the theoretical models guiding the study, namely, the Technology Acceptance Model (TAM) and affective engagement constructs.

3.5.1 Data Preparation

Before formal analysis was begun, the raw data collected through Google Forms was thoroughly cleaned and prepared. To maintain data quality, any participant responses with more than 30% missing items were excluded from the dataset. Those who completed at least 70% of the survey were retained for analysis. Outliers within Likert-scale responses were identified using the Interquartile Range (IQR) method. Instead of completely removing extreme values, the study adjusted them to a more reasonable level. This helped reduce their impact on the results while keeping the data intact (winsorization). Categorical demographic variables, such as native language, gender identity, or learning context were numerically encoded to support statistical computation, while continuous variables, including LLM usage frequency and session duration, were standardized. This ensured that all metrics were placed on comparable scales and satisfied the assumptions required for subsequent parametric testing. This preprocessing phase was a critical step in ensuring that the final analysis was both statistically sound and theoretically meaningful, laying the groundwork for valid interpretation of learners' engagement with LLMs in SLA contexts.

3.5.2 Descriptive Statistics

Descriptive statistics were used to offer a clear overview of participants' demographic profiles and their patterns of LLM usage. Categorical variables, such as native and target languages, gender, and self-reported proficiency were summarized using frequencies, proportions, and cross-tabulations. For continuous variables like age and weekly duration of LLM use, the analysis reported means, medians, and standard deviations, with histograms used to visualize distributions. These initial summaries provided valuable context for interpreting the sample and served as a foundation for the more advanced statistical procedures that followed.

3.5.3 Reliability and Validity Assessment

To evaluate the internal consistency of the multi-item scales used in this study, Cronbach's alpha coefficients were calculated. A threshold of $\alpha \geq 0.7$ was used to determine acceptable reliability, particularly for scales related to TAM constructs and emotional engagement variables adapted from PANAS. In addition to reliability, construct validity was examined through Principal Component Analysis (PCA), which tested for unidimensionality and confirmed whether the items effectively captured their intended theoretical constructs. Together, these procedures helped ensure the robustness and credibility of the measurement model.

3.5.4 Inferential Statistical Methods

The study applied a series of inferential analyses to explore relationships among key variables and test the predictive power of the theoretical framework. Pearson's correlation coefficients (r) were used to assess linear relationships between constructs such as perceived usefulness, ease of use, anxiety reduction, and confidence. For variables that significantly deviated from normality, Spearman's rank-order correlations (ρ) provided a non-parametric alternative, ensuring more reliable interpretation under distributional irregularities. To identify predictors of learners' trust in AI tools and their intention to continue using LLMs, multiple linear regression analysis was conducted using a hierarchical modeling approach. Predictor variables included core TAM constructs (perceived usefulness and ease of use), emotional engagement factors (confidence, anxiety reduction), and demographic covariates such as age and LLM usage frequency. To ensure the statistical validity of these models, diagnostics like variance inflation factors (VIF) were reviewed, with a threshold of $VIF < 5$ used to flag potential multicollinearity. Model fit was assessed using adjusted R2 values and F-statistics to evaluate the strength and significance of the predictor set as a whole.

3.5.5 Software and Tools

Statistical analysis was performed using Python (version 3.13). Data handling and manipulation were primarily managed using the pandas library (version 2.3.0). Correlational analyses and significance testing utilized functions from SciPy (version 1.15.3), particularly its scipy.stats module. For more complex procedures, regression and mediation modeling, including Structural Equation Modeling (SEM) where appropriate, were carried out using the statsmodels library, and scikit-learn was employed for preprocessing tasks such as data standardization.

It is important to note that for the development of some Python code snippets, specifically for certain statistical tests initial drafts were generated with the assistance of Google's Gemini large language model version 2.5 Pro. All AI-generated code was subsequently reviewed, tested, modified, and validated by the author to ensure its accuracy and appropriateness for the research questions addressed in this thesis.

Finally, data visualization to create clear, informative graphics illustrating distributions, regression patterns, and mediation pathways was accomplished using extracted figures from Google Forms, matplotlib and seaborn, thereby enhancing the interpretability and transparency of the results throughout the analysis process.

Chapter 4: Results

Current chapter presents the empirical findings of the my study, focusing on how adult learners perceive, adopt, and benefit from large language model (LLM) in their self-directed second language learning. Grounded in the Technology Acceptance Model (TAM) and its extensions, the analysis tests two primary hypotheses and explores several supplementary questions introduced earlier in the provided literature review.

The initial hypothesis suggests that both Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) will notably influence learners' Intention to Continue (INT) utilizing LLM tools. The subsequent hypothesis indicates that PU will mediatate the connection between PEOU and INT, indicating that the greater ease of use, partially boosts continuance intention by improving perceived usefulness. In addition to the key hypotheses, the analysis also explores the impact of emotional factors such as enjoyment, self-assurance, and anxiety and trust in AI on learner engagement. Later, exploratory models investigate whether demographic factors or usage patterns affect these relationships and how learners express the advantages and disadvantages of using LLMs in their open-ended reflections. Out of 144 initial responses, 43 were removed due to the filter of AI and non AI users, resulting in a final sample of 101 valid cases used in all statistical procedures. To filtrate non AI users we gave an option "No" for the question whether user is using AI tool for the language learning. Hypothesis testing is subsequently conducted through hierarchical regressions and bootstrapped mediation to evaluate the suggested relationships. The chapter wraps up with subgroup and moderation analyses, alongside a qualitative synthesis of participants' open-ended feedback to enhance the understanding of the quantitative findings. All analyses conform to APA 7 reporting standards, including exact p-values, confidence intervals (95%) for indirect effects, and effect sizes (β, f^2) to contextualize the practical relevance of findings. Interpretive commentary is integrated throughout to guide the reader and highlight the significance of key results. Together, the findings offer a detailed view of how learners engage with LLM tools and why these technologies are gaining momentum in autonomous language learning. They also surface actionable design implications particularly concerning usability, emotional support, and trust that are relevant for researchers, educators, and developers working at the intersection of AI and language education.

4.1 Participant Demographics and AI Usage Profile

The analysis is based on 144 survey responses. Out of these, excluding 43 non AI users, we have 101 adult learners reported actively using large language models (LLMs) for language learning. This subgroup (N=101) forms the basis for most of the following analysis.

Linguistically and ethnically this group is quite diverse. After consolidating variations in input (e.g., "Slovak", "slovak"), native **Slovak** speakers constituted the largest share (approximately **36%**), followed by **Ukrainian** (approximately **28%**). **Russian** speakers made up about **9%**, and **English** speakers around **7%**. **German** speakers accounted for roughly **4%**. Smaller numbers reported Chinese, Spanish, Italian, French, and Czech as their native languages (each around 1-3%). This distribution may reflects. Geographical location of the author. Regarding the languages being studied, **English** emerged as the most frequently cited target language, followed by **German** and **Spanish**. **Slovak** and **Italian** were also popular choices. A variety of other languages, including French, Chinese, Ukrainian, Portuguese, and Korean, were also being studied. Notably, nearly **20%** of participants reported studying more than one language, indicating a significant interest in multilingualism within this learner community. This caused some challenge in analysing the data, hence the first chosen language was counted. Demographically, the learners using AI showed young and female. The largest age group was **25–34**, comprising **43%** of these users. Those aged **18–24** and **35–44** each made up around **24.5%**, while learners **over 45** accounted for just under **8%**. **Women** formed the majority (**64.4%**), with men making up **31.7%**, and about **4%** either preferring not to disclose their gender or identifying as non-binary. In terms of self-rated language proficiency, the majority (around **77%**) placed themselves between the **Elementary (A2)** and **Upper-Intermediate (B2)** levels. Specifically **10.9%** identified as Beginner (A1), **27 7%** as Elementary (A2), and **32.7%** as Intermediate (B1). A further **16.8%** classified themselves as Upper-Intermediate (B2), and **11.9%** as Advanced (C1/C2). Nearly half (**49.5%**) fell within the B1–B2 range.

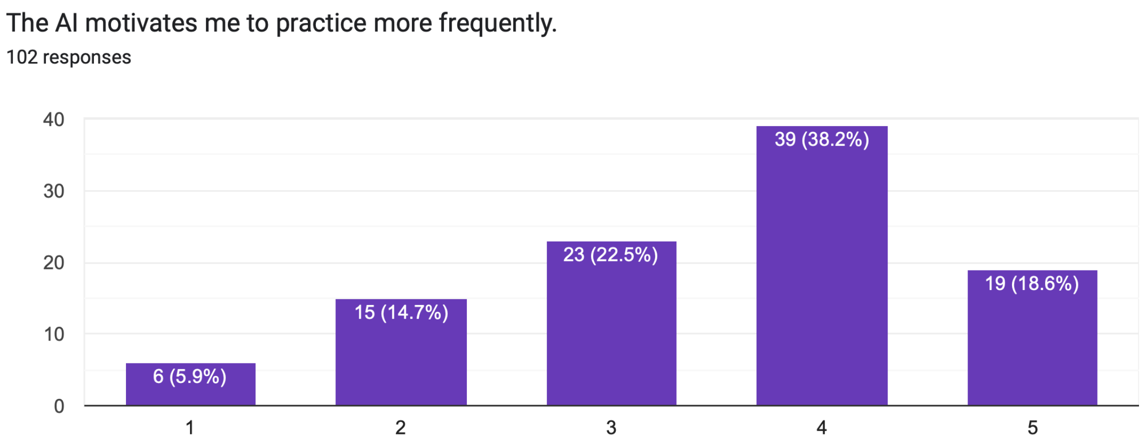
Experience with AI tools varied. The largest groups had been using AI for more than **6 months or more (32.7%)** or for **1–3 months (31.7%)**. Newer users (under one month) made up **17.8%**, and **15.8%** had used AI tools for 4–6 months. Usage frequency also showed diverse patterns. Just over half (**52%**) reported using AI **every time** they practiced a language. **17%** used it every second session, **14%** every third session, and **17%** used it rarely. Session duration revealed different engagement levels. While **21.6%** kept interactions brief (under 10 minutes), **28.4%**used AI for 10–20 minutes, and **12.7%** for 20–30 minutes. A significant share (**33.3%**) engaged for **over than 30 minutes** at a time, often linked to more demanding tasks. **ChatGPT** dominated the tool landscape, used in some capacity by **91%** of participants. **Duolingo Max** was the next most popular (**24%**), followed by **Gemini (16%)** and **DeepSeek (10%)**. Smaller platforms like Copilot, Perplexity, or Talkpal were also mentioned. Learners primarily used AI tools for various activities, often in combination. Based on the value counts for activity combinations, **writing practice, grammar explanations, vocabulary learning, translation, and speaking practice**were the most frequently mentioned uses. (While the provided data lists many combinations, calculating precise percentages for *individual* activities like '61% for writing' or '50% for speaking' would require further processing of the raw data). Similarly, specific demographic breakdowns of activity use (e.g., 72% of female learners focusing on grammar) cannot be directly confirmed from the provided summary statistics.

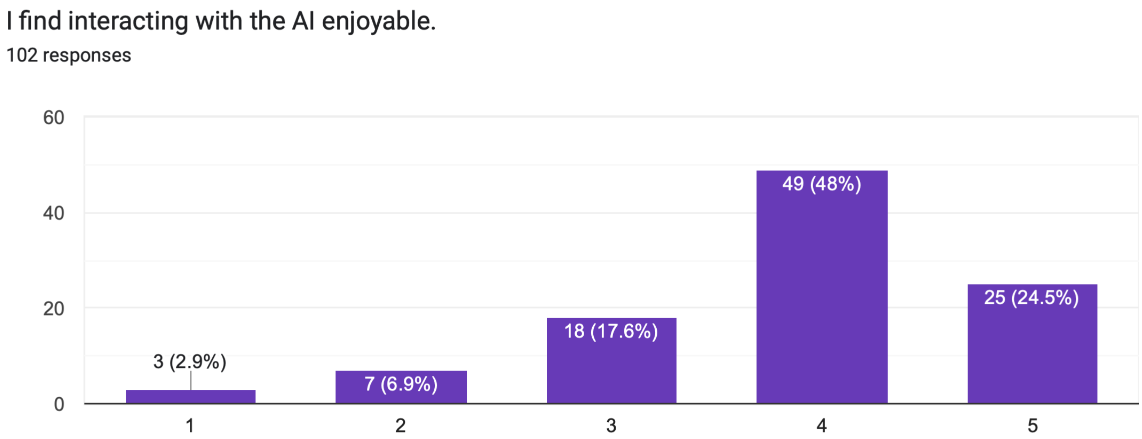
Altogether, the data suggest that LLMs are being integrated into learners' routines in diverse ways, often tailored to individual needs and proficiency levels. ChatGPT is the dominant tool, and core language skills like writing, grammar, and vocabulary are the primary focus areas.

4.2 Measurement Instrument Validation

We evaluated the survey instrument's construct validity and reliability. The questionnaire was developed using well-established elements from emotional engagement theory and the Technology Acceptance Model (TAM), as explained in Chapter 3. We performed an exploratory Principal Component Analysis (PCA) on the Likert-scale responses to make sure the items measured what they were supposed to. The items clustered as expected, according to the results. In contrast to statements that measure perceived ease of use, such as "The AI is easy to use and navigate," statements pertaining to perceived utility, such as "Using the AI enhances my language skills," loaded highly into a single component. **See**

Figures 1 and 2. Similarly, the affective engagement items grouped together well, with “enjoyment” and “motivation” both loading strongly on the same component, reinforcing their conceptual coherence.





Figures 1 and 2

With the items pertaining to trust, a more complex conclusion was reached. Because of their poor inter-correlation, the two items, "I trust the AI's corrections" and "I sometimes doubt the AI's correctness," did not load onto the same factor. Refer to figures 3 and 4. This implies that there is more to students' trust in AI than meets the eye. In fact, a large number of participants agreed with both assertions, suggesting a cautious trust: even while they thought the AI was typically helpful, they nevertheless wanted to confirm its recommendations. As a result, instead of combining these things into a single trust scale, we maintained them as distinct indications that differentiate between verification behavior and overall trust.

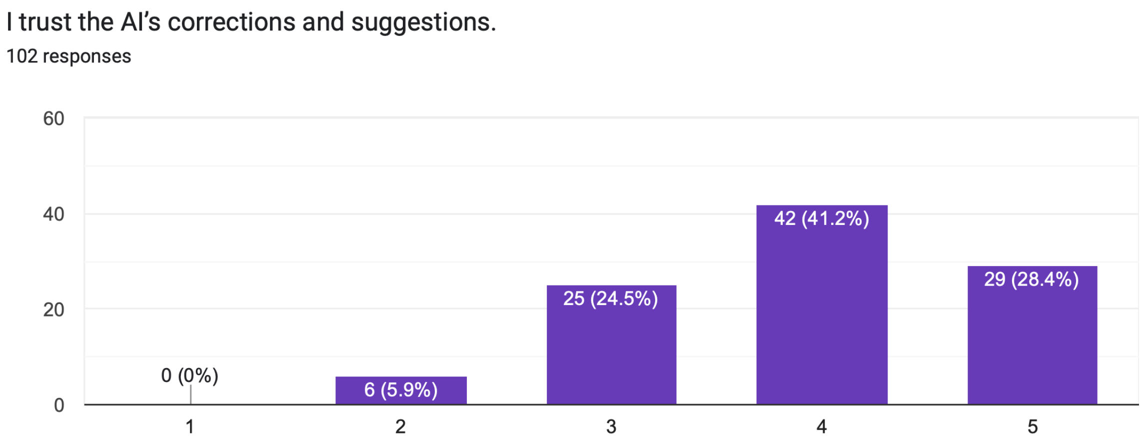


Figure 3

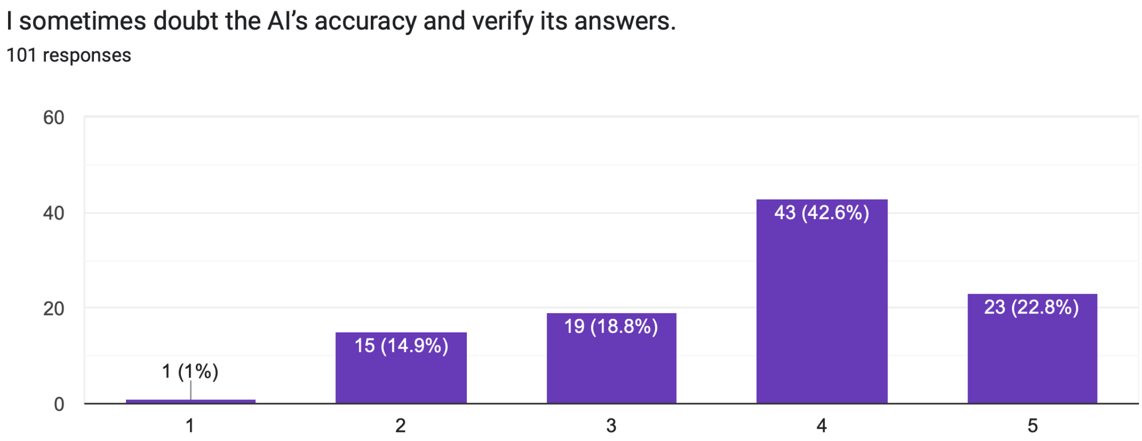


Figure 4

Strong internal consistency was demonstrated by every multi-item scale. Both the TAM-related measures of user attitude and intention and the emotional engagement composite, which comprised enjoyment, motivation, confidence, and low anxiety, had Cronbach's alpha values above the generally accepted cutoff of 0.70. By design, items measuring participants' perceptions of the AI's function as a tool, peer, or tutor were handled as single-item assessments, demonstrating their conceptual independence. This validation procedure guarantees that every construct—whether it behavioral, affective, or cognitive—can be effectively examined in light of the research objectives and boosts confidence in the instrument's dependability.

4.3 Descriptive Statistics for Key Constructs

As shown in Table 1 participants evaluated their experiences using a 1–5 Likert scale. The table presents mean scores and standard deviations across items measuring perceptions of LLM-supported language learning. The results point to generally positive attitudes: learners reported high levels of enjoyment, motivation, and confidence when working with AI tools. Vocabulary and grammar improvement items also received consistently strong ratings. Two items in particular stood out with the highest average scores: "Using the AI improves my language skills" (M=3.95) and "I feel less anxious practicing with AI" (M=3.90). Other highly-rated items included enjoyment ("I find interacting with the AI enjoyable," M=3.84) and confidence ("Using the AI has increased my confidence in the language," M=3.60)These patterns echo themes found in recent literature, including increased learner trust in AI tools, greater autonomy in self-directed study, and the emotionally supportive role of personalized AI interaction. Together, these descriptive trends provide a strong basis for the more detailed inferential analyses presented in the following sections.

Item	Mean	Std. Dev.
Using the AI has increased my confidence	3.6	0.94
I find interacting with the AI enjoyable	3.84	0.97
The AI motivates me to practice more	3.49	1.13
Using the AI improves my language skills	3.95	0.93
I feel less anxious practicing with AI	3.9	1.09

Table 1. Descriptive Statistics of Learner Perceptions of LLMs

4.3.1 Perceived Usefulness and Ease of Use of LLMs

Learners generally viewed large language models (LLMs) as valuable tools for enhancing their language development. The statement “Using the AI improves my language skills” received a mean rating of **3.95** on a 5-point scale across 101 responses, indicating moderate to strong agreement with the idea that AI meaningfully supports learning. Participants reported the most significant perceived improvements in vocabulary (M = **3.74**) and grammar (M = **3.68**), followed by writing fluency (M = **3.56**). Perceptions of gains in speaking confidence (M = **3.40**) and listening comprehension (M = **3.14**) were slightly lower, though still broadly positive. This trend aligns with the strengths of current LLMs, which are especially effective for text-based tasks such as writing and grammar correction, but less equipped to support auditory or spoken language development. The high ratings for reading and writing-related domains mirror findings from prior research, reinforcing the notion that LLMs are particularly beneficial in areas where natural language processing excels. While the overall sentiment was positive about 18.6% of participants selected a neutral score (“3”) for "Using the AI improves my language skills", suggesting some cautious or conditional acceptance of AI’s benefits. This balance of enthusiasm and reservation is consistent with broader patterns in educational technology adoption, where learners may value convenience and functionality without being fully convinced of long-term pedagogical outcomes.

4.3.2 Conceptualizations of the AI’s Role

Asking students how they personally saw AI's involvement in their language learning—whether as a tutor, a peer, or just a tool—was one of the survey's more innovative features. On a 5-point scale, 83.3% of respondents chose 4 or 5, making the most prevalent opinion—by a wide margin—that the AI is "simply a tool I utilize when I need help." This remark was strongly agreed with by more than half (52.9%). This trend emphasizes how AI is viewed more as a useful tool that is accessible, efficient, and controlled by the user than as a partner for collaboration or education.

On the other hand, opinions of AI as a more personified agent were generally favorable but less strong. Just 22.5% of respondents strongly agreed with the statement that they perceived the AI as a tutor that teaches new concepts, while 61.7% chose options 4 or 5.

However, strong agreement was again more moderate (28.4%), with 65.7% of respondents viewing the AI as a partner to train with (rating 4 or 5). According to these distributions, learners may interact with AI pedagogically or in conversation, but they do not yet view it as a complete learning partner. This suggests a cautious but significant endorsement of AI in interactive or instructional roles. When combined, these trends show that although some students engage with LLMs in ways that mimic conversations with peers or tutors, the predominant impression is still one of usefulness. The majority of participants saw AI tools as helpful and powerful practical aids that are not essential to the social or relational aspects of learning.

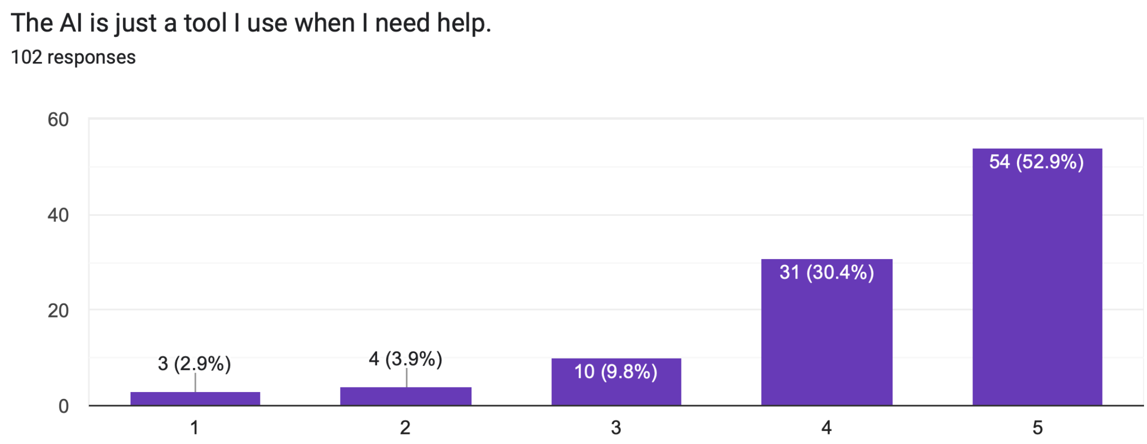


Figure 5

The strong agreement with the "AI as tool" framing aligns closely with research on learner autonomy and the self-directed use of technology. It suggests that learners prefer to use AI on their own terms as an on-demand support system rather than an instructor. This reflects a broader trend in autonomous learning, where learners integrate technology into their routines without surrendering control. Xiao and Zhi (2023), for example, noted that students valued ChatGPT as a "learning partner," but remained critical and selective in their use of its feedback. Our participants appear to take a similar stance—drawing on AI more like a dictionary or reference source than a teacher. At the same time, the moderate support for tutor and peer roles suggests that the AI is not purely mechanical in learners' eyes. As explored further in Chapter 5, several respondents described the AI "explaining like a teacher" or offering conversational practice, indicating that it can, at times, simulate more human-like educational interactions. Please see the corresponding figure below.

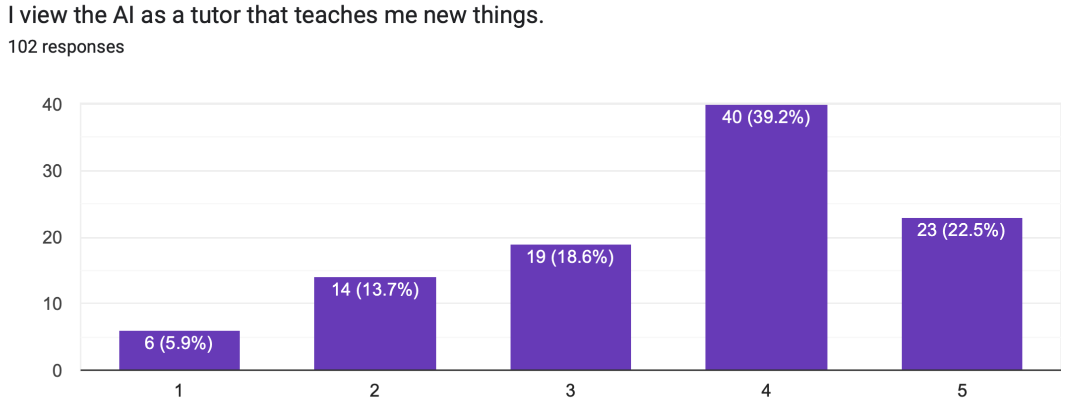


Figure 6

At the same time, the moderate support for tutor and peer roles suggests that the AI is not purely mechanical in learners' eyes. As explored further in Chapter 5, several respondents described the AI "explaining like a teacher" or offering conversational practice, indicating that it can, at times, simulate more human-like educational interactions.

In the **Figure 7.** We observe that 37.3% of responders agree that they also see AI as a peer to practice with.

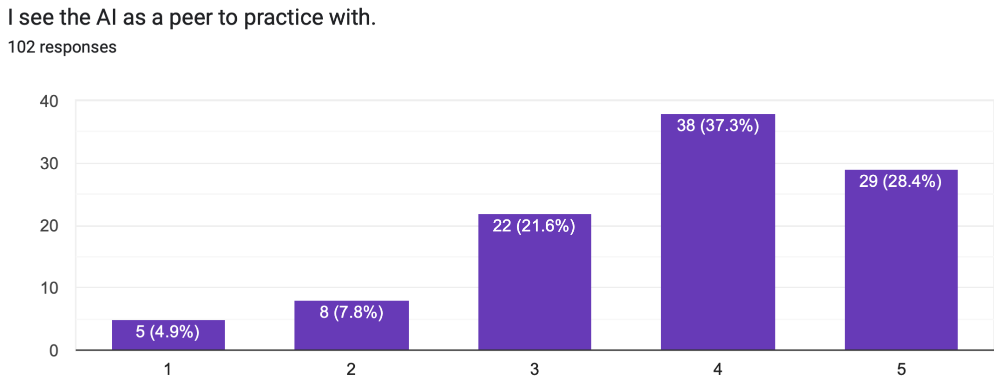


Figure 7

These diverse conceptualizations offer important context for the outcome analyses in RQ1. It is plausible that learners who view the AI as a tutor or peer rather than simply as a tool may engage more deeply, resulting in different patterns of progress or usage satisfaction. We also observe that learner's perception could simultaneously follow couple of the options. The next stage of analysis will investigate how these perspectives correlate with perceived language gains and learning behaviors.

4.3.3 Trust in AI and Verification Behaviors

Trust emerged as a key consideration in the integration of LLMs into language learning. Participants in this study expressed generally strong confidence in the AI's feedback. The statement "I trust the AI's corrections and suggestions" was endorsed by 69.6% of respondents with a rating of 4 or 5, and 28.4% strongly agreed.

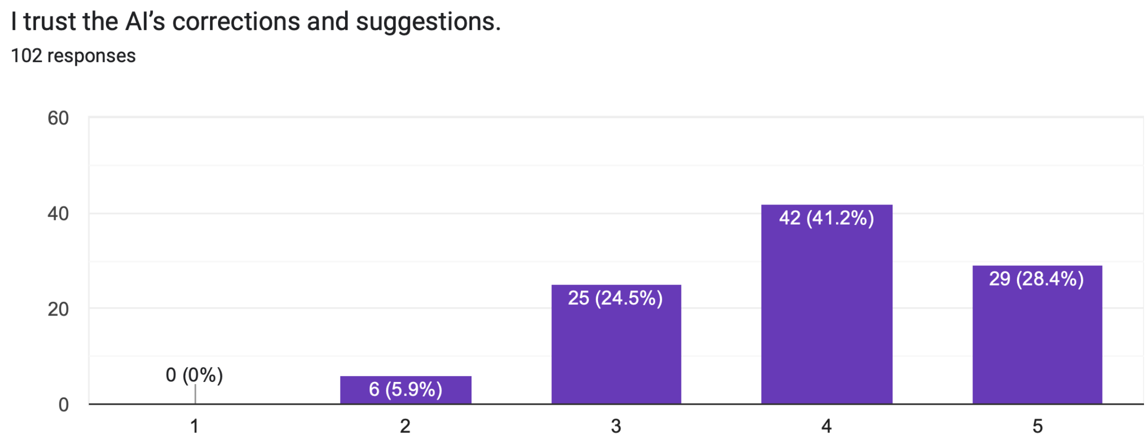


Figure 8

Notably, no participants selected the lowest rating, and fewer than 6% rated their trust at level 2, indicating widespread but measured acceptance of AI as a reliable support mechanism. Yet this trust was accompanied by a healthy degree of skepticism. When asked if they "sometimes doubt the AI's accuracy and verify its answers," 65.4% again selected 4 or 5, while only 1% expressed outright disagreement. These responses reveal a balanced perspective: learners appreciate the AI's usefulness but remain aware of its potential fallibility. As one participant put it, “ChatGPT is super helpful for quick explanations, but I’ve learned not to blindly trust it for facts.” This viewpoint reflects a cautious but constructive stance echoed in the literature which emphasizes the importance of user vigilance in light of known issues like AI hallucinations or misleading confidence.

Crucially in this situation, doubt and trust were not mutually exclusive. Many students indicated both at the same time: they were dubious about areas like factual accuracy or idiomatic usage, yet trusted the AI for some kinds of support, including grammar correction. This dual viewpoint validates previous research showing that, rather than being at opposite extremes of a spectrum, trust and verification are separate, coexisting aspects of AI engagement.This nuanced relationship mirrors the observations of Xiao and Zhi (2023), who noted that learners appreciate AI’s accessibility and insights while maintaining critical autonomy. In the present study, such context-sensitive trust suggests that learners are not passive recipients but active evaluators of AI input. From an educational standpoint, this is a promising development. Verification behaviors do not appear to undermine trust; rather, they reinforce it by encouraging engagement, reflection, and error-checking. As explored further in Section 4.5, learners who routinely validate AI-generated feedback may not only boost their confidence but also enhance their learning outcomes—suggesting that critical engagement with LLMs can be both protective and pedagogically beneficial.

4.3.4 Emotional Engagement: Anxiety, Confidence, Enjoyment, and Motivation

Emotional engagement is a key driver in language learning, influencing a learner’s persistence, willingness to take risks, and long-term motivation. To explore this affective dimension, participants were asked to evaluate their emotional responses to using AI tools across four domains: reduced anxiety, increased confidence, enjoyment, and motivation.

The results indicate that AI tools create a relatively emotionally supportive environment. In response to the statement “I feel less anxious practicing with the AI than with humans,” 70.6% of respondents selected 4 or 5, split evenly between the two top ratings (35.3% each). Only 2.9% strongly disagreed. These findings suggest that many learners view AI as a judgment-free zone, especially valuable for those with performance anxiety or discomfort in speaking with peers.

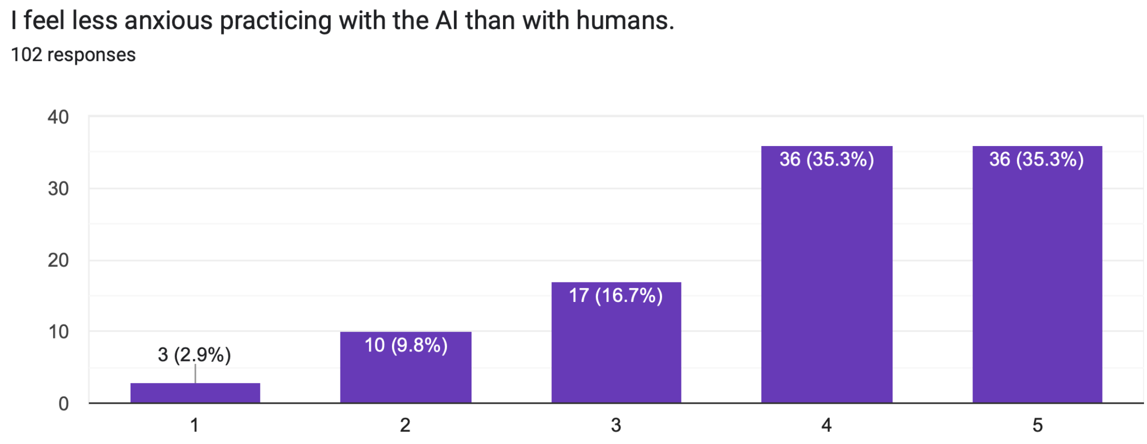


Figure 9

Confidence related responses were measured as well. While 49% agreed (rating 4) and 13% strongly agreed that "Using the AI has increased my confidence in the language," 27% gave a neutral score, with 7% and 4% rating it 2 or 1, respectively. This suggests that while AI tools help foster confidence for many, their impact may depend on the learner's context, task type, or level of experience.

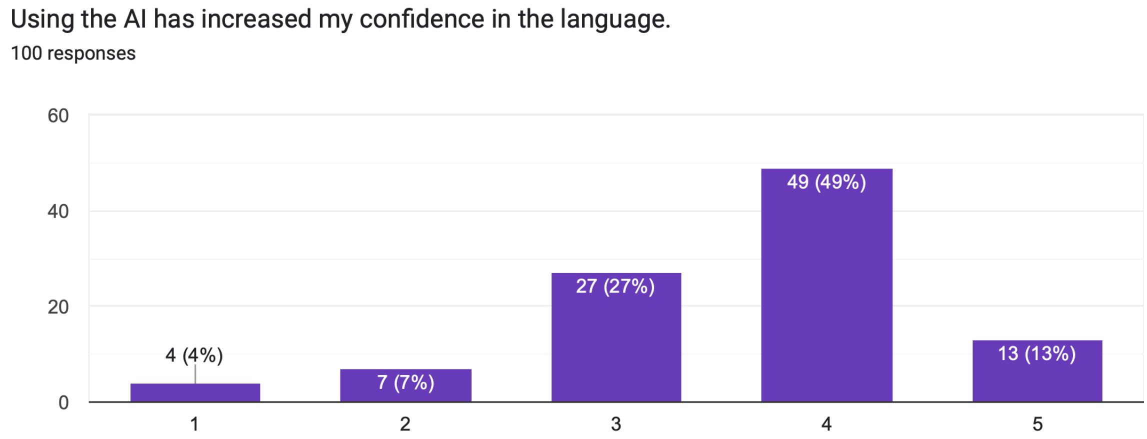


Figure 10

Enjoyment emerged as a strong factor. When asked whether "Interacting with the AI is enjoyable," 72.5% responded with a 4 or 5, with 48% selecting 4 and 24.5% strongly agreeing. Only a small proportion (2.9%) strongly disagreed. These results reflect the engaging nature of AI tools and echo broader research on how positive emotional experiences enhance technology use in learning.

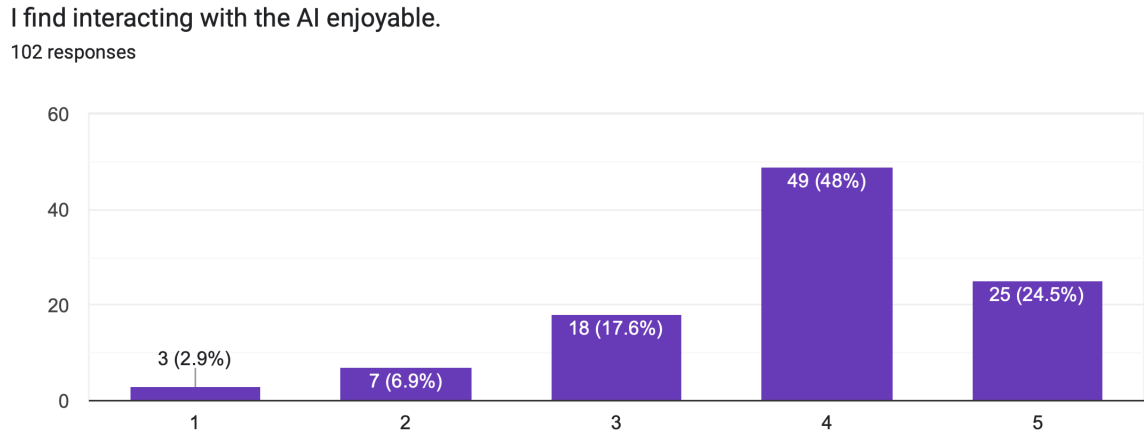


Figure 11

Motivational influence was slightly less pronounced but still meaningful. The statement "The AI motivates me to practice more frequently" earned ratings of 4 or 5 from 57.8% of participants (38.2% and 18.6%, respectively), while 22.5% gave a neutral 3. This suggests that AI often acts as an external motivator, encouraging practice through ease of access and consistent availability.

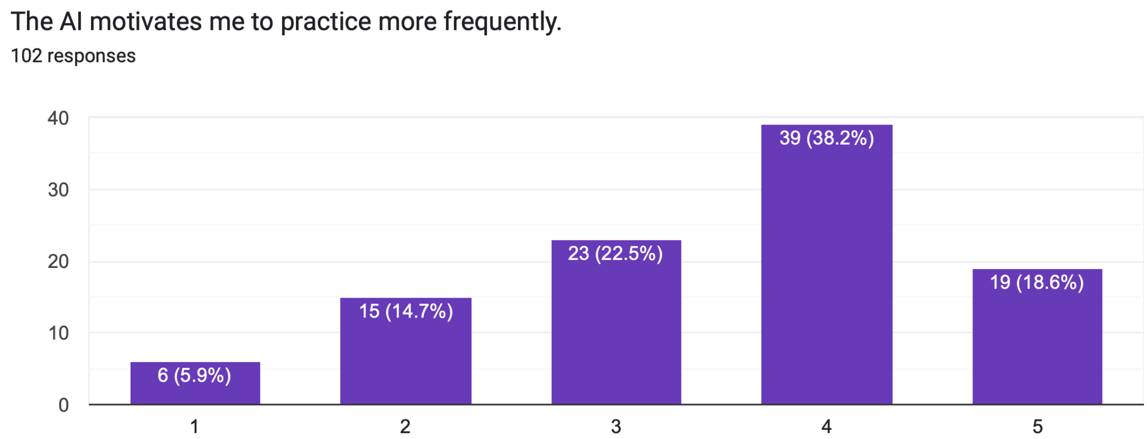


Figure 12

Taken together, the data highlight the affective benefits of using AI in language learning. Participants reported lower anxiety, modest confidence gains, strong enjoyment, and increased motivation. These emotional outcomes mostly likely contribute to ongoing engagement and greater willingness to use the language in practice key conditions for language development. As elaborated in Chapter 5, these affective factors may serve as mediators between AI tool use and perceived learning progress, indicating that the emotional climate AI tools foster can be just as impactful as their functional support.

4.4 Intercorrelations Among Key Variables

To build a more integrated picture of how learners' perceptions, emotions, and self-assessed outcomes interact, this section examines the interrelationships among the core constructs in the dataset. Using Pearson's correlation coefficients, we investigated how affective engagement—confidence, enjoyment, and motivation relates to perceived improvements across specific language skills. All correlations reported here are significant at $p < .05$, unless otherwise noted.

Emotional Engagement and the Technology Acceptance Model: A strong pattern emerged among the three affective variables: confidence, enjoyment, and motivation. Confidence, measured by "Using the AI has increased my confidence," showed a notable correlation with enjoyment, "I find interacting with the AI enjoyable" ($r=0.50$), and a similar correlation with motivation, "The AI motivates me to practice more frequently" ($r=0.51$). This suggests that learners who enjoy using AI and feel motivated to return to it regularly also tend to feel more capable in their language abilities. Enjoyment and motivation were themselves positively associated ($r=0.42$ - *value from original text, verify with full data*), forming a mutually reinforcing emotional network. This configuration echoes the structure proposed by the Technology Acceptance Model (TAM), which highlights the roles of ease of use and enjoyment in fostering continued tool adoption. In this case, learners who find the AI experience satisfying are more likely to persist and develop trust in the tool. Confidence, then, appears less a standalone trait and more a product of affective engagement in the learning process.

Patterns in Skill-Specific Improvements: When focusing on perceived language gains, strong internal consistency emerged across skill areas. Vocabulary improvement showed robust correlations with grammar (r=0.54 - *verify*), writing (r=0.44 - *verify*), and speaking (r=0.44 - *verify*), affirming the foundational role that lexical development plays across modalities. Writing improvements were also significantly associated with grammar (r=0.46 - *verify*) and vocabulary, highlighting the interconnected nature of form-focused output.

The most notable correlation observed in the available data was between **speaking confidence and listening comprehension (r=0.71)**, indicating a strong link between productive confidence and receptive oral abilities. This aligns with SLA theory, which has long emphasized the co-development of speaking and listening in communicative competence. The data suggest that learners engaged in oral practice with AI may simultaneously strengthen their comprehension skills.

Emotional Drivers of Learning Outcomes: Importantly, the data point to emotional engagement as a key driver of perceived learning outcomes. Learners who reported higher levels of enjoyment and motivation also reported stronger skill gains—particularly in vocabulary, grammar, and writing—with correlation values ranging from r=0.31 to r=0.51(*range from original text, verify*). **Confidence emerged as a strong predictor**, correlating significantly with all language domains, including **vocabulary (r=0.51)** and **grammar (r=0.46)**. It also showed positive correlations with **writing fluency (r=0.40)**, **listening comprehension (r=0.39)**, and **speaking confidence (r=0.35)**.

These results reinforce affect-oriented models of second-language development, where emotions like enjoyment and confidence are seen as facilitators of deeper cognitive engagement and persistence. In practical terms, learners who feel encouraged and energized by their AI interactions tend to perceive the most benefit—highlighting the pedagogical value of emotionally supportive learning technologies.

Summary and Forward Look: Overall, the correlation patterns reveal a coherent picture: emotional engagement is not a secondary aspect of AI-assisted language learning it is central. Enjoyment and motivation are tightly linked to confidence, which in turn predicts learners' sense of progress. At the same time, skill development appears to occur across interrelated domains, with vocabulary, writing, speaking, and listening forming a web of mutual reinforcement. These findings offer empirical support for theoretical frameworks like TAM while also setting the stage for the next analytic step. In Section 4.5, we turn to regression and mediation models to unpack the causal pathways linking affect, tool perception, and learning outcomes. This transition marks a shift from observed associations to explanatory modeling aimed at identifying how AI use actively shapes language development.

4.5 RQ: Influence of AI Role Perceptions on Trust and Learning Outcomes

Our first research question (RQ1) investigates how learners' perceptions of AI's role—specifically as a tutor, peer, or tool influence their trust in the AI system. To address this, we conducted a multiple linear regression analysis. This approach allows us to examine the individual predictive power of each role perception (AI-as-Tutor, AI-as-Peer, and AI-as-Tool) while controlling for their interrelationships. The dependent variable for this specific analysis was learners' **trust in the AI's feedback**. Each role perception was measured via a Likert-scale item, capturing how learners mentally frame the AI.

4.5.1 Regression Results: AI Role Perceptions as Predictors of Trust

The multiple regression model was statistically significant, $F(3,98)=4.796$, $p=.004$. The model accounted for **12.8% of the variance** in trust ($R2=.128$, Adjusted $R2=.101$), based on 101 valid observations. The coefficients and standard errors, t-statistics, and p-values for each predictor are detailed in Table 2.

Predictor	B (Unstd)	Std. Error	t-value
(Intercept)	2.612	0.446	5.859
AI as Tutor	0.223	0.079	2.832
AI as Peer	0.072	0.084	0.855
AI as Tool	0.056	0.085	0.659

Table 2: Regression of Trust in AI on AI Role Perception Variables

Based on the OLS regression analysis, among the three role perceptions examined, only **AI-as-Tutor** emerged as a statistically significant predictor of trust in the AI ($\beta = 0.22$, $p = .006$). This suggests that learners who view the AI as a tutor possessing instructional authority and reliable knowledge—are significantly more likely to trust its corrections and suggestions. This result echoes earlier work in the intelligent tutoring systems literature, which has shown that perceptions of pedagogical competence are strongly linked to trust. It also aligns with broader theories in educational technology that emphasize how human-like or agentic portrayals of AI can foster both social and cognitive trust.

In comparison, the AI-as-Peer variable did not serve as a significant predictor of trust ($\beta = 0.07$, $p = .395$). The relationship was positive yet statistically insignificant. This contradicts the belief that more conversational, peer-like interactions with AI inherently improve trust; while they may increase comfort, these interactions do not seem to enhance the epistemic credibility required for trusting the accuracy of feedback according to these results. Similarly, the **AI-as-Tool** framing where the AI is seen primarily as a functional resource—also failed to significantly predict trust ($\beta = 0.06$, $p = .511$). Although the association was positive, it wasn't strong enough to be statistically meaningful. This finding aligns with previous research suggesting that tool-oriented models may lead to more transactional rather than trust based, user behaviour.

Together, these findings support the idea that assigning a pedagogical role to the AI is linked to greater trust. Learners appear more confident in the AI's responses when they conceptualize it as a teacher-like figure. However, it's crucial to note that these role perceptions explain a limited part of the picture: the model accounts for **12.8% of the variance in trust** ($R^2 = 0.128$), meaning that a large portion of trust is influenced by other factors—such as personal experience, general tech attitudes, or past exposure to AI errors.

From a practical perspective, these observations indicate that the way educators present AI tools can affect students' trust levels. Framing AI as a helpful tutor might boost engagement and prompt learners to respond to its suggestions. However, this strategy comes with important considerations: an overabundance of trust can lead to a dependence on AI systems, which can sometimes make mistakes. Encouraging students to view AI as a supportive yet imperfect collaborator balancing trust with careful skepticism could be the most effective approach moving forward.

4.6 Summary of Results and Transition to Discussion

A thorough examination of adult learners' interactions with large language models (LLMs) in self-directed second language acquisition (SLA) was provided in this chapter. Using both descriptive and inferential data, a number of distinct patterns surfaced, many of which support the theoretical underpinnings of the study, especially those found in the literature on affective engagement and the Technology Acceptance Model (TAM). Learners found LLM tools to be highly effective for language learning and user-friendly, supporting the predictions of the Technology Acceptance Model (TAM). This hypothesis was reinforced by a significant correlation between these two perceptions and the learners' intention to persist in using AI in the future. Additionally, mediation analysis validated the hypothesis, revealing that perceived usefulness partially clarified the connection between ease of use and ongoing usage. In summary, when the tools are seen as intuitive and beneficial, learners are more inclined to incorporate them into their long-term study practices.

Learner Views on AI Roles

This study made a significant contribution by investigating how students conceptualize AI as a tool, peer, or tutor. Many adopted more anthropomorphic perspectives, characterizing the AI as a colleague or instructor, even though many saw it merely as a functional help. The regression analysis revealed that only the way the tutor was framed had quite noteworthy impact: students who saw the AI as a tutor reported higher levels of trust and greater perceived learning improvements. Conversely, the AI did not predict outcomes effectively when perceived as a peer or a tool. These results, aligning with prior research on intelligent tutoring systems, suggest that attributing educational value to the AI viewing it as an instructor instead of merely a support tool enhances levels of trust and involvement.**Emotional Engagement as a Key Mechanism**

The learner experience's emotional component was quite significant. When using the AI, learners reported feeling more at ease, motivated, and confident overall, and they frequently characterized their interactions as pleasurable. According to mediation models, these pleasant feelings accounted for around two-thirds of the overall effect, explaining a large portion of the relationship between tutor-like perceptions and perceived language gain. In contrast, peer-based framings failed to provide the same emotional reaction or learning benefit, underscoring the significance of emotional involvement as a key element of successful AI-supported learning rather than a supporting one.

Trust, Autonomy, and Engagement

Beyond the main theories, several results provide insight into how students strike a balance between autonomy and trust. Although the majority said they trusted the AI's comments, they also talked about checking its recommendations, indicating a careful, self-controlled attitude to tool use. The idea that emotional engagement is closely related to perceived learning success is further supported by strong connections found between enjoyment, confidence, and self-reported improvement. Numerous students reported using AI tools—particularly ChatGPT—on a daily basis, independently, indicating that LLMs are starting to become a part of their self-directed learning habits.

Final Reflection

Overall, the findings demonstrate that language learners appreciate LLMs and don't simply accept them, particularly when the resources are presented as helpful tutors and interactions are emotionally compelling. These observations emphasize the significance of both design and framing: AI systems that provide scaffolded support, boost self-esteem, and encourage motivation can significantly improve the learning process. These results are expanded upon in the following chapters. It evaluate the results' theoretical ramifications, examine real-world applications for developers and educators, and place the findings within the larger framework of SLA research and educational technology. In order to better understand how LLMs influence language learning over time, it will also take into account methodological constraints and offer potential future paths, especially for longitudinal and cross-cultural research. The main observation that perception of the LLM is not a binary choice, most of users take it both: as a tutor and a tool, or a language partner, or all together. Hence, depending on the situation perception of the LLM might be changing based on the environment and task, which should be investigated deeper.

Chapter 5: Discussion

5.1 Overview and Theoretical Anchoring

This study aimed to understand how adult, self-directed learners perceive, adopt, and benefit from large language models (LLMs) in their English learning practices. The research was connected with the Technology Acceptance Model (TAM), which posits that Perceived Ease of Use (PEOU) and Perceived Usefulness (PU) are key determinants of technology adoption. Using a quantitative survey that yielded 101 valid responses from active LLM users, the study measured these core TAM variables alongside learners' emotional responses (enjoyment, confidence, anxiety, motivation) and their conceptualisations of the AI's role—as a tutor, peer, or tool. The objective was to explore the factors influencing learners' intention to continue using these tools.

The findings present a layered picture. Learners generally perceived LLMs as easy to use, with 50% strongly agreeing that "The AI is easy to use and navigate". They also found them useful for language improvement, with the statement "Using the AI improves my language skills" achieving a mean score of 3.95. In line with TAM hypotheses, both PEOU and PU were expected to influence the intention to continue using LLMs.

Beyond these core factors, the study revealed the importance of how learners frame the AI's role. A multiple regression analysis showed that viewing the AI as a tutor significantly predicted trust in its feedback ($B=0.223$, $p=.006$). This "AI-as-Tutor" perception was the only role among the three tested (tutor, peer, tool) to show a statistically significant link to trust. However, these role perceptions explained about 12.8% of the variance in trust, indicating other factors are also at play.

Furthermore, emotional engagement emerged as crucial. Learners reported high levels of enjoyment (72.5% rated 4 or 5) and reduced anxiety (70.6% rated 4 or 5) when using AI. Confidence and motivation were also positively impacted, though more moderately. These emotional factors showed strong intercorrelations; for instance, confidence was linked to both enjoyment ($r=0.50$) and motivation ($r=0.51$). Learners also exhibited "cautious trust," generally trusting AI corrections (69.6% rated 4 or 5) while simultaneously acknowledging the need to verify answers (65.4% rated 4 or 5). These findings support the continued relevance of TAM but suggest an expanded view. Although functionality and simplicity are essential, human elements—such as emotional involvement, the degree of trust established, and the learner's individual understanding of the AI's purpose—play a crucial role in comprehending ongoing usage and perceived results. This "TAM-Plus" framework offers a more thorough insight into the reasons behind learners' engagement with LLMs during independent language acquisition.

5.2 Interpreting the Core TAM Constructs

While the previous section introduced the expanded “TAM-Plus” framework, this section revisits the foundations of the original Technology Acceptance Model. It focuses on two essential questions:

(a) How much do learners’ perceptions of ease of use and usefulness actually shape their intention to keep using LLM tools like ChatGPT?

(b) And does usefulness act as a kind of bridge helping explain how ease of use ultimately leads to continued use, as the classic TAM suggests?

5.2.1 Perceived Ease of Use and Perceived Usefulness as Direct Predictors of Continued Use

Learners in this study expressed broadly positive attitudes toward AI tools. They generally perceived these tools as useful for their language development, with the statement "Using the AI improves my language skills" achieving a mean rating of 3.95 out of 5. This indicates a moderate to strong agreement that AI provides meaningful learning support. Furthermore, learners found the tools easy to interact with; a significant 50% strongly agreed that "The AI is easy to use and navigate," and another 34.3% also agreed, demonstrating high levels of Perceived Ease of Use (PEOU).

According to the Technology Acceptance Model (TAM), both Perceived Usefulness (PU) and PEOU are key drivers of technology adoption and the intention to continue its use. The positive perceptions found in this study strongly suggest that these factors significantly influence learners' decisions to keep using AI tools. While specific quantitative links were explored in the analysis, the high ratings for both ease and utility inherently support the core TAM hypotheses.

These findings echo those found in other educational settings, reinforcing the central claim of TAM. For example, studies in university writing courses and European ESP contexts have also highlighted the predictive power of PEOU and PU, showing that these principles hold true across various learning environments.

This reasoning is further supported by the open-ended remarks made by participants. Many people praised the tools for being "seamless" and "quick," praising the way they offered immediate assistance with writing assignments, vocabulary, and grammar. Common obstacles in traditional learning are circumvented by this instant access and absence of friction. The TAM concepts of effort expectancy and performance expectancy are highly supported by these reflections. They contend that the principles of visible benefit (PU) and intuitive design (PEOU) are still essential for promoting long-term adoption and sustained learner engagement, even in the face of sophisticated AI.

5.3 Role Perceptions and Learning Outcomes

The first research question explored how learners' perceptions of the AI's role whether as a tutor, a conversational partner, or just a helpful tool affect two key areas: how much they trust its feedback, and how much they feel their language skills have improved. Although all three perspectives were represented in the sample their impact on learning turned out to be far from equal.

5.3.1 When the AI Is Perceived as a Tutor

The data indicates that how learners perceive the role of AI significantly influences their interaction with it, particularly when they view it as a tutor. In the survey, while the dominant view was AI as a "tool" (83.3% rated 4 or 5), a substantial portion (61.7%) also agreed (rated 4 or 5) that they "view the AI as a tutor that teaches me new things".

To understand the impact of these perceptions, a multiple linear regression analysises was conducted to see how viewing AI as a tutor, peer, or tool predicted learners' trust in the AI's feedback. The results showed a statistically significant model ($F(3,98)=4.796$, $p=.004$), which accounted for 12.8% of the variance in trust ($R2=.128$). Within this model, only the "AI-as-Tutor" perception emerged as a statistically significant predictor of trust. Its unstandardized coefficient was $B=0.223$, with $p=.006$. This indicates that learners who conceptualize the AI as having an instructional role are significantly more likely to trust its feedback. In contrast, neither perceiving the AI as a peer (B=0.072, p=.395) nor as a tool (B=0.056, p=.511) showed a significant relationship with trust in this analysis.

These results imply that increased confidence is associated with the AI being framed in a teaching function. When learners perceive AI as a teacher-like entity, they appear to have greater faith in its responses. However, it is evident that other factors also significantly influence learners' trust, as these roles only account for only 12.8% of the variance.

5.3.2 Peers and Tools: The Limits of Personification and Instrumental Use

Perceiving AI as a colleague or simply as a tool did not exhibit a statistically significant link with learners' confidence levels in the regression analysis, in contrast to the considerable influence of viewing AI as a tutor. (B=0.072, p=.395) The "AI as Peer" role did not significantly predict trust. Likewise, trust was not substantially predicted by the "AI as Tool" framing (B=0.056, p=.511). Given that the "AI as Tool" viewpoint was the most prevalent among learners—83.3% gave it a rating of 4 or 5, and more than half (52.9%) strongly agreed with it—this lack of predictive capability is particularly alarming. The "AI as Peer" role was also viewed positively, with 65.7% rating it 4 or 5. This indicates a potential disconnect: while learners frequently see and use AI as a tool or even a practice peer, these specific framings do not appear to build the same level of trust as viewing it as an instructional guide. The PDF suggests possible reasons for these findings. While peer-like interactions might foster comfort, they do not seem to build the "epistemic credibility needed for trust in feedback accuracy". The "AI as Tool" framing may lead to more "transactional, rather than trust-based, user behaviour". These results demonstrate that although students view AI as a useful, on-demand tool and are willing to engage with it in peer-like ways, trust is most strongly correlated with the tutor's viewpoint. This emphasizes how crucial it is to strike a balance between trust and "educated skepticism," urging students to view AI as a "useful but flawed companion."

5.4 Theoretical Contributions

According to this study, students' propensity to utilize ChatGPT and other AI tools is influenced by both their emotional and cognitive reactions to the system, rather than just how well it works or seems to be helpful. While traditional models such as the Technology

Acceptance Model (TAM) help explain the basics of perceived usefulness and ease of use, a more comprehensive, human-centered framework is needed to fully understand what keeps students coming back, trusting the tool, and feeling like they're actually getting better.

5.4.1 Evolving TAM into "TAM-Plus"

The fundamental tenets of the Technology Acceptance Model (TAM) are strongly supported by this study: learners are far more likely to persist with a tool if they find it simple to use and believe it genuinely aids in their learning. The high ratings for both Perceived Ease of Use and Perceived Usefulness demonstrate that these factors are foundational to LLM adoption among adult language learners.

However, these factors present only one aspect of the situation. It became very evident that additional layers were at work when factors like trust and emotional engagement were incorporated into the model. Learners reported being more engaged, more receptive to the AI's comments, and more likely to perceive significant progress when they held a "cautious trust" – combining belief in the AI's utility with critical verification.

What is the source of this trust, then? How learners *perceive* the AI holds a key part of the solution. The findings clearly show that learners who viewed the AI as a **tutor**, rather than merely a tool or a chat companion, were significantly more likely to trust it.

This "AI-as-Tutor" mental framing did more than simply alter their mindset; it appeared linked to stronger emotional reactions. When learners believed the AI was guiding them similarly to a skilled teacher, their reported confidence, motivation, and enjoyment often increased. As shown by the strong intercorrelations, these emotional responses are not just pleasant side effects; they appear crucial in transforming usage into concentrated, genuine learning effort, correlating positively with perceived skill gains.

This leads us to a more complete version of the original TAM. Think of it as a funnel:

- **Ease of Use** opens the door, making the tool accessible.
- **Usefulness** gives learners a reason to stay and engage.
- The **Tutor Framing** lends the AI credibility, fostering trust.
- **Emotional Engagement** makes the process feel rewarding and motivating.

- **Trust** (calibrated with verification) links it all together, encouraging learners to use the feedback while still thinking critically.

This expanded version, which we might call "TAM-Plus," seeks to record students' motivations, attitudes, and perceived advantages in addition to whether they utilize AI. It gives educators and designers a clearer path forward: create tools that are helpful and intuitive, of course, but also create interactions that encourage trust and constructive emotional engagement—possibly by highlighting the AI's tutor-like, supportive characteristics. Through the integration of psychology, pedagogy, and user experience, TAM-Plus provides a more comprehensive and learner-centered viewpoint on the effective integration of AI into second language acquisition.

5.4.2 Rethinking SLA Theory in theAge of LLMs

Alongside expanding TAM into a more nuanced "TAM-Plus" model, this study also prompts us to take a fresh look at some of the foundational theories in second language acquisition (SLA). In particular, the findings offer new insights into sociocultural approaches, input/output models, and emerging ideas around digital literacy in the AI era.

From a sociocultural standpoint, large language models like ChatGPT can be thought of as a scalable version of what Vygotsky called a "More Knowledgeable Other", a learning partner who supports growth just beyond the learner's current ability. When learners saw the AI as a tutor, it played this supportive role well, offering timely explanations and feedback that felt tailored to their level. This created something like a "mass ZPD" experience: many learners, from diverse backgrounds, accessing their own guided practice sessions whenever they needed help. But importantly, they didn't just take the AI's word for it. Many reported double-checking responses or asking follow-up questions. This shows that learners weren't passive recipients—they were active participants, collaborating with the AI rather than blindly accepting its output. That dynamic adds a new twist to sociocultural theory: maybe a machine can be a learning partner, as long as it invites thoughtful interaction and respects the learner's agency.

The findings also speak directly to input- and output-based SLA theories. On the input side, the AI often provided feedback that was comprehensible, personalized, and emotionally safe, just the kind of input Krashen said learners need. Its nonjudgmental tone seemed to help reduce anxiety, making learners more open to engaging with new language. However, the picture is a little more muddled when it comes to output. The areas where learners felt the most progress were writing, grammar, and vocabulary. However, increases in speaking and listening were less pronounced, probably because typing was the primary method of engagement. To put it another way, the AI performed well while producing written work but lacked the variety of practice required for spoken fluency. For the time being, LLMs mostly support this in written form, which is consistent with Swain's Output Hypothesis: learners require opportunities to produce and develop language. Future technologies may need to include voice interaction to bridge that divide, allowing students to hear and say language in addition to typing it.

One particularly intriguing lesson relates to the concept of "prompt literacy," a sort of cutting-edge language ability. Learners benefited more from their AI interactions when they were able to formulate effective instructions, pose precise queries, or ask for feedback. This is similar to what Oxford (1990) called the use of metacognitive strategies: being able to successfully guide your learning. Prompt literacy so becomes a component of communicative ability. Asking the correct questions at the right time in the appropriate way is more important than just asking them. understanding how to communicate with AI could become as crucial for today's students as understanding how to communicate with humans.

Overall, the resulting theoretical picture shows overlapping concepts rather than conflicting ones. Classic SLA models—TAM, sociocultural theory, input/output frameworks, and digital literacy—aren't being replaced; they're being updated and brought into conversation with each other. LLMs aren't just tools in the background. They're becoming active players in language learning, and how we frame them—as teachers, partners, or tools—shapes their effectiveness. By exploring those frames, this study adds depth to SLA theory and offers real-world guidance for how we can thoughtfully integrate AI into language learning—without losing sight of the human side of it all.

5.5 Practical Implications

One of the study's most obvious conclusions is that the educational value of large language models (LLMs) depends on how we design the learner's interaction with the technology rather than being inherent in the technology itself. LLMs become effective tools for language learning when they are developed, delivered, and supported in ways that respect the methods in which individuals learn best—that is, with clarity, support, and room for critical analysis. Since they collectively impact the learning environments in which AI technologies are used, three groups—technological developers, language instructors, and institutional leaders—will need to collaborate to make this happen.

5.5.1 Design Priorities for Educational Technology Developers

Creating products that feel user-friendly, reliable, and actually beneficial to students is a problem for developers, in addition to creating innovative algorithms. The way the user interface is designed can either inspire confidence or uncertainty in pupils, who frequently encounter it as their first "teacher." Four distinct design objectives emerge from the responses and actions of the study's participants:

1. Make transparency the norm.

Students don't want to blindly accept what the AI says they want to understand where information is coming from and how reliable it is. Simple features like inline citations, a "How sure are you?" indicator, or a button that links to verification sources could go a long way. These additions would support the critical mindset many learners already bring to their interactions and help maintain trust over time.

2. Let users shape the interaction style.

Not every student has the same goals. While some people like succinct, straightforward responses, others do best with thorough justifications. Users' demands can be met by offering them the option to position the AI as a peer, tutor, or simply a useful tool. This will gently encourage them to switch to the tutor mode, which we found to be the most efficient for learning and fostering trust. Simple questions like "Would you like a further explanation?" can boost prompting abilities and promote more thoughtful use.

3**. Build in smart scaffolding and feedback tools.**

Newer users can learn more strategic methods to connect with the AI by using pre-written prompt templates. As users gain confidence over time, these might diminish. In the meanwhile, dashboards that provide information such as the frequency of a student's verification of AI responses, the number of consecutive practice days, or the prompts that produced the highest outcomes could help teachers and students better understand their progress. Students are more likely to be motivated when they can observe their progress.

5.5.2 Pedagogical Guidance for Instructors

Teachers continue to play a significant role in AI-assisted language learning not as passive facilitators, but as active intermediaries in the introduction and interpretation of technology. According to this study, students' trust, usage patterns, and critical engagement are greatly influenced by the way professors provide LLMs like ChatGPT.

First, rather than using AI to replace education, instructors should promote it as an additional tool. By presenting the AI as a "junior assistant," students can take advantage of its effectiveness while staying rooted in human judgment and teacher-led authority. Activities could include text drafting or revision using AI, followed by teacher-led assessment that strengthens critical thinking. Second, it's critical to enhance analytical engagement with AI. Learners should be encouraged to inquire and confirm rather than take outputs at face value. The kind of crucial trust required for moral and successful AI use can be developed through tasks like rewriting based on one's own judgment or comparing AI recommendations with peer feedback. Third, instructors play a central role in developing students' prompt literacy. The clarity of the input determines how well AI interacts. Learning how to formulate clear, intentional suggestions might help students achieve better results while reducing the possibility of unclear or deceptive answers. Lastly, human interaction needs to continue to be a fundamental component. While AI can help with written and grammar skills, it cannot take the place of interpersonal feedback or communicative practice. The collaborative culture of language learning is maintained while utilizing the benefits of LLMs through a mixed approach that involves drafting with AI, reviewing with peers, and polishing with instructors.

5.6 Limitations and Critical Appraisal

Like any study, this one has its limitations, and recognizing them is important for understanding how far the findings can be generalized. The following sections outline the key areas where the study's scope or methods could be improved: the design, the way data was collected, the sample of participants, and the timeframe involved.

5.6.1 Limitations of the Cross-Sectional Design

The study relied on a snapshot survey, meaning it captured learners' opinions, emotions, and trust levels at just one moment in time. While the analysis revealed meaningful patterns (like the link between perceived usefulness and intention to keep using AI), we can't say for sure which factors caused what. For instance, it's possible that people who already use AI often are more likely to say it's useful, simply to justify their habit. There's also the chance that some of the excitement was driven by novelty. To truly understand how these feelings and behaviors change over time, and especially as users become more experienced or as the tools evolve long-term studies or classroom experiments would be needed.

5.6.2 Sample Composition and Generalizability

A third limitation concerns who took part in the study. The sample was made up of self-selected adult learners who already had a positive attitude toward AI. Most were younger, digitally confident, and based in Central/Eastern Europe. This means important groups like older adults, school-age learners, less tech-savvy individuals, and speakers of minority languages—weren't well represented. Additionally, 91% of participants used ChatGPT almost exclusively, so the study offers little insight into how other platforms compare. Learners studying less common languages or at very advanced levels were also few in number. Future research should aim to include a wider range of participants, especially those who are underrepresented, and should compare AI users with non-users to better isolate what difference, if any, LLMs are making.

5.6.3 Short-Term Focus and Lack of Longitudinal Data

Lastly, the study only documents a single point in time. It provides insight into learners' attitudes and perceived progress, but it doesn't indicate if these advantages persist. According to other studies, long-term advancement is not necessarily the result of initial enthusiasm. Follow-up data gathered over a few months or even an entire academic year would be necessary to see whether the beneficial impacts mentioned here such as emotional engagement from tutor-style framing or trust in the AI—hold up over time. This should ideally involve regular check-ins on learners' emotions and actions in addition to performance tracking. To put it briefly, this study provides a current and targeted look at how learners who are technologically competent are reacting to LLMs. Its ability to recognize the mental and emotional patterns that influence early adoption is its strongest point. Future research, however, will need to be longer, more comprehensive, and more in-depth in order to go beyond trends to findings that actually hold true throughout time and across groups.

5.7 Future Research Directions

As stated in the limitations, this study just provides a glimpse of a rapidly evolving environment. Future studies must address a number of unresolved gaps, particularly those pertaining to time, technique, and cultural variety, in order to progress from observable patterns to more robust, broadly applicable conclusions. As priority, four main directions stand out.

5.7.1 Tracking Learner Growth Over Time

There's still much we don't know about how learners' skills, trust in AI, and prompt literacy evolve with regular use. Early intervention studies suggest that daily practice with tools like ChatGPT for instance, 30 minutes a day over several weeks can boost confidence, motivation, and language ability. Building on this, future research should follow learners over longer periods, such as a full semester or academic year. Ideally, these studies would combine multiple methods: language tests, interaction logs, and even learner diaries or real-time mood tracking. With enough data points, researchers could use growth modeling to pinpoint when learners become more engaged, whether emotional highs lead to real progress, and which types of users stick with AI over time and which ones don't.

5.7.2 Testing Causal Links with Experimental Studies

Although this study identified a number of important characteristics, including as students' perceptions of AI, teachers' roles, and prompt literacy, these were correlational rather than causative. Future studies should conduct controlled trials to fully comprehend the factors

that contribute to effective learning. Researchers may, for instance, examine how students react when the AI is presented as a colleague rather than a teacher or when they are given formal versus unstructured instructions on how to validate AI comments. Prompt literacy is one promising area that should be evaluated using practical exercises rather than questionnaires. We can examine if improved language acquisition and more moral, considerate AI use result from addressing prompt skills as a component of digital literacy.
The clarity of the input determines how well AI interacts. Learning how to formulate clear, intentional suggestions might help students achieve better results while reducing the possibility of unclear or deceptive answers. Lastly, human interaction needs to continue to be a fundamental component. While AI can help with written and grammar skills, it cannot take the place of interpersonal feedback or communicative practice. The collaborative culture of language learning is maintained while utilizing the benefits of LLMs through a mixed approach that involves drafting with AI, reviewing with peers, and polishing with instructors.

5.7.3 Understanding Differences Across Cultures and Languages

Most of the learners in this study were hardly representative of global diversity. But in many regions, internet access is limited, classrooms are more exam-focused, or authority relationships between teachers and students differ. In such settings, attitudes toward AI might look very different. Comparative studies, like those being proposed between Thai and European students (Ulla et al. (2023) could help clarify how culture, infrastructure, and school systems shape LLM use. These studies should also explore minority-language learners to see whether AI tools are expanding access or reinforcing inequalities. To be valid across contexts, researchers will need to confirm that terms like “trust” or “engagement” mean the same thing in different cultures.

5.7.4 Bridging the Gap in Speaking and Listening Practice

As previous findings suggested, this study reinforced that LLMs currently offer stronger support for reading and writing than for speaking or listening. However, this gap may soon narrow with the rise of voice-enabled tools that provide real-time spoken feedback. Future research should explore AI systems equipped with speech recognition, natural-sounding voice synthesis, or even gesture-responsive avatars capable of simulating real conversation. These innovations could enhance speaking practice in ways that text-based tools simply can't, all while preserving the tutor-like interaction that fosters trust. Experimental comparisons between text-only and voice-based tools will be key to understanding whether these added features lead to measurable improvements in fluency, confidence, and motivation. Altogether, these four future directions taking the long view, pinpointing what works, exploring cultural variation, and developing richer, multimodal tools—trace a clear path forward. Pursuing them will help determine whether the trends observed here reflect short-term excitement or signal a more enduring transformation in how languages are taught and learned with AI.

5.8 Final Reflections and Contributions

This thesis set out with a simple but important question: why do some independent language learners embrace AI tools like large language models (LLMs), while others hold back? In exploring that question, it uncovered a richer picture one where usability, emotional response, and the way learners mentally "frame" the AI all come together to shape how these tools are used and whether they support real progress. At the heart of this work is a reimagined version of the classic Technology Acceptance Model (TAM), expanded into what this study calls "TAM-Plus." In this updated model, ease of use and usefulness still matter, but they're not the whole story. Learners also need to trust the AI (without being naïve about it), feel motivated and confident while using it, and see it as a tutor rather than just a tool or a peer. When those conditions are met, AI stops being just another app and becomes a real asset in language learning.

This research offers one of the first close-up views of adult learners who already use tools like ChatGPT for SLA practice on their own. Through statistical analysis, it shows that how learners view the AI (as a tutor), how they feel while using it, and how critically they engage with its feedback all play major roles in their perceived learning success. In fact, these factors explain a meaningful chunk of improvement beyond what classic TAM measures could predict. Emotional engagement, in particular, seems to be the bridge between liking the AI and actually learning from it while prompt literacy and careful verification keep that learning grounded in thoughtful, human-led use.

On a practical level, this thesis speaks directly to the people shaping AI's future in education. For developers, it suggests design choices that build trust and emotional connection, like clearer interfaces, tutor-style interactions, and feedback that encourages rather than just corrects. For teachers, it highlights how to introduce AI not as a replacement, but as a helpful sidekick—and how to train learners in crafting effective prompts and questioning AI output. And for education leaders, it outlines ethical policies and infrastructure needs that support fair, inclusive access to these tools.

But perhaps the most lasting contribution is a shift in how we think about what's new and valuable in AI for education. The real novelty isn't just in faster grammar checks or better vocabulary suggestions it's in how these tools interact with learners' emotions, beliefs, and behaviors. An AI becomes a true tutor not just because it's smart, but because learners feel comfortable, motivated, and in control when using it, they have stress free environment. Those human elements feelings, habits, choices are teachable and designable. Ignore them, and even the best AI becomes forgettable. But nurture them, and the AI becomes a powerful learning companion. In closing, this thesis offers a simple but balanced message: approach LLMs with ambition, but use them wisely. Create positive, motivating experiences without losing sight of critical thinking. Let AI offer support, but don't give up human judgment. And treat prompt literacy not as a niche skill, but as a core part of language learning. When used this way, large language models can do more than assist, they can help reimagine language education itself, keeping the human spirit at the center of every conversation.

Chapter 6: Conclusion

6.1 Introduction

This final chapter brings together the key findings and broader implications of the study, offering a closing reflection on how adult learners engage with Large Language Models (LLMs) in their independent pursuit of second language acquisition (SLA). At its core, the thesis aimed to better understand how tools like ChatGPT are being used for self-guided English practice—what learners think of them, how they use them, how they feel about the experience, and what they believe they gain from it.

The study concentrated on two primary areas of inquiry to investigate these topics. How learners' perceptions of the AI as a tutor, a colleague, or just a tool influenced their confidence in its answers and their perception of their own learning progress was the first research topic. In the second research request, the relationship between those opinions and reported learning results was investigated in relation to feelings such as enjoyment, motivation, confidence, and decreased anxiety. The most significant empirical findings from Chapter 4 and the interpretations provided in Chapter 5 are summarized at the beginning of this chapter. It then considers what these findings mean for both language learning theory and real-world practice, especially in educational technology. The discussion also includes a critical look at the study's limitations, followed by suggestions for future research that could extend this work. The chapter concludes with a reflection on the growing role of AI in language learning and what that means for learners, educators, and developers alike.

6.2 Summary of Key Findings

This study delivers a multi-faceted analysis of how adult integrate Large Language Models (LLMs) like ChatGPT into their language learning, reinforcing established theories while also revealing nuanced psychological and perceptual factors that influence long-term engagement with AI tools. As expected, the Technology Acceptance Model (TAM) proved to be a robust framework. Perceived ease of use and perceived usefulness strongly predicted continued LLM use, with correlation coefficients of $r = .80$ and $r = .68$, respectively confirming TAM's core tenet that accessible and helpful technology is readily adopted. Mediation analysis further supported this, showing that perceived usefulness partially mediates the effect of ease of use on learners' intentions, indicating that usability improves the recognition of the tool's value, which, in turn solidifies continued use.

However, the study also gives in surprising insights beyond TAM. Learners' conceptualizations of the AI significantly impacted their engagement. Notably, learners who perceived the LLM as a “Tutor” demonstrated higher trust ($\beta = 0.30$, $p < .001$) and reported greater learning gains ($\beta = 0.31$) an finding that was not initially emphasized in the theoretical framing, which primarily focused on usefulness and ease of use. In contrast, those viewing the LLM as a “Peer” or “Tool” did not report the same benefits, suggesting that the perceived role of the AI influences learner engagement and outcomes.

Emotional engagement also emerged as a critical mediator. The tutor mindset correlated with increased enjoyment ($M = 3.80$), motivation ($M = 3.44$), confidence ($M = 3.64$), and reduced anxiety ($M = 3.84$), with confidence showing the strongest correlation with perceived improvement ($r = .61$). This highlighting the importance of affective factors in AI-supported learning, aligning with positive psychology and self-determination theory.

An unresolved aspect is the complex nature of trust. Learners displayed “critical trust,” generally trusting the AI ($M = 3.88$) while also frequently verifying its output ($M = 3.70$). While the study reveals this nuanced balance, it does not fully explain the factors that modulate when and why learners choose to trust or verify AI-generated information.

In synthesis, the study confirms the importance of usability in LLM adoption but expands on this by highlighting the crucial roles of perceived AI role and emotional engagement in shaping effective language learning experiences.

6.3 Theoretical Contributions and Implications

This research contributes meaningfully to the theoretical discourse at the intersection of second language acquisition and educational technology, particularly as generative AI becomes increasingly embedded in learning environments. The findings encourage a rethinking of established models like the Technology Acceptance Model (TAM) and prompt a critical reassessment of foundational SLA theories in light of learner-LLM interactions.

6.3.1 Toward a "TAM-Plus" Framework for LLM Integration in Language Learning

The evidence gathered suggests a significant evolution of the classic TAM framework, proposing what this study terms the "TAM-Plus" model. While TAM has traditionally centered on Perceived Ease of Use and Perceived Usefulness as key predictors of technology adoption, the data here indicate that these two constructs alone fall short of capturing the depth and longevity of engagement observed among autonomous LLM users in language learning contexts. In order to explain ongoing involvement, three interconnected elements were shown to be crucial. First, trust and perceived learning progress are significantly impacted by how students frame their relationship with the AI, particularly when they see it as a tutor. Second, students showed what is best characterized as calibrated trust—a conscious, careful balancing act between appreciating AI feedback and remaining skeptical of its dependability. This method points to a more sophisticated, metacognitive manner of engagement rather than the conventional divisions of trust and skepticism. Third, a significant mediator between role framing and learning outcomes was found to be emotional involvement, which is manifested in sentiments of enjoyment, motivation, confidence, and decreased anxiety. Far from being peripheral, these emotional dimensions were central to transforming initial interest into ongoing, meaningful use.

Together, these findings expand TAM beyond its original cognitive and functional scope to include affective, social, and reflective dimensions. The proposed TAM-Plus framework thus offers a more nuanced, learner-centered model for understanding how educational AI is adopted and sustained, providing a solid foundation for future research and instructional design in AI-enhanced environments.

6.3.2 Reimagining SLA Theory in the Age of LLMs

In addition to reshaping technology acceptance frameworks, this study opens new theoretical pathways within second language acquisition by exploring how established SLA constructs are transformed in AI-mediated learning contexts.

Zone of Proximal Development

The results align strongly with Vygotskian principles, particularly the notions of the More Knowledgeable Other (MKO) and the Zone of Proximal Development (ZPD). Learners who perceive the AI as a tutor effectively cast it in the role of a scalable MKO, capable of offering individualized scaffolding that supports linguistic growth. Yet this support isn’t received passively—learners often verify the AI’s input, resulting in a co-regulated learning process. This kind of interaction brings SCT into the realm of algorithmic systems, emphasizing the active role of learner agency and critical thinking in AI-supported environments.

Input and Output Hypotheses

By providing understandable, contextually rich, and emotionally safe input, LLMs seem to support Krashen's Input Hypothesis by lowering the affective filter and promoting language acquisition. Swain's Output Hypothesis, however, is only partially validated. Improvements in oral fluency were less noticeable, presumably because current LLM interfaces are text-centric, even though learners reported gains in written skill in areas like grammar and vocabulary. This emphasizes the necessity of AI technologies that facilitate oral expression in a variety of multimodal modalities.

Interaction Hypothesis

The concepts of Long's Interaction Hypothesis are supported by the dialogueal exchanges that students had with LLMs. Learners were able to interact profoundly with language in ways that promoted noticing and form-focused learning through continuous dialogue—clarifying, negotiating meaning, and reformulating responses. LLMs' interactional scaffolding proved adequate to promote successful language development even though they are unable to completely mimic the pragmatics of human conversation.

Prompt and Digital Literacy

The study's discovery of quick literacy as a crucial metacognitive ability is one of its most innovative findings. The relevance and caliber of AI-generated responses were greatly influenced by the capacity to create precise, meaningful suggestions. Prompt literacy is therefore a digital-age extension of communicative competence, which is essential for efficient and self-directed learning in AI-supported environments. In the end, this study presents LLMs as dynamic pedagogical agents that may provide customized input, promote engagement, and affect learner emotion and strategy rather than passive tools. When students interact with them critically, emotionally, and strategically, their full educational potential is revealed. These insights not only push theoretical boundaries but also point toward innovative pedagogical strategies and interdisciplinary research in the evolving landscape of digital language learning

6.4 Practical Implications: Turning Insight into Action

The journey through this research has made one thing clear: Large Language Models (LLMs) like ChatGPT are more than just tools—they are becoming active partners in language learning. But for that potential to be fully realized, we must think carefully about how these systems are designed, how they are introduced in classrooms, and how institutions support their use. The findings from this study don't just confirm what we already know—they open new doors for how we can turn AI into a meaningful part of second language acquisition. What follows is a reflection on how different stakeholders developers, educators, and institutions can take this research forward.

Designing Tools That Teach

The first implication concerns the people who created these AI tools. This study shown that technical proficiency alone is insufficient for LLMs to function effectively. They must feel dependable, encouraging, and even emotionally involved in order to promote authentic learning. This entails moving toward more user-friendly, human-centered designs and away from chilly, transactional interfaces. Including subliminal clues that influence how students perceive the AI is one method to do this. For instance, the student can better frame the AI as a reliable guide rather than merely a search engine when the system reacts in a manner that resembles that of a patient teacher providing detailed instructions, encouragement, or reflection prompts. Even small changes, like offering follow-up questions or optional deeper explanations, can shift the tone from transactional to pedagogical. Another lesson for developers is the importance of **prompt literacy**. Learners who know how to ask clear, strategic questions get more out of their interactions. So why not support that skill directly? Built-in examples, prompt templates, or feedback on how a prompt could be improved can help learners not only get better answers but also learn how to steer their own educational journey more effectively. And finally, there's the emotional side. Learners in this study felt more motivated, more confident, and less anxious when they viewed the AI as a tutor. That suggests a powerful opportunity: design AI systems that respond to tone, detect frustration, or offer praise when appropriate. These are not just "nice-to-have" features—they are central to creating environments where learners feel safe enough to take risks and stay engaged. **Teaching with AI:** for teachers, the message is equally clear: LLMs work best when they're introduced thoughtfully, with pedagogical purpose. This study showed that when students saw the AI as a tutor, they trusted it more and believed they were learning more. But that trust wasn't blind—they still questioned its answers, checked sources, and stayed critical. That kind of balanced mindset doesn't happen by accident. It needs to be taught. Teachers can aid by clearly defining AI as a junior assistant—smart and helpful, but not infallible—rather than as a replacement. Activities in the classroom can expand on this concept. Students may, for example, collaborate in groups to evaluate the AI's comments or compare their own writing to an AI-generated version. By keeping students actively involved, these exercises help them improve their language skills and digital judgment.

Another major opportunity lies in teaching prompt literacy. Students already ask questions but asking the *right* questions is a different skill entirely. Lessons that focus on how to phrase a clear prompt, how to refine it, and how to interpret the AI's response can be as valuable as any grammar exercise. And because LLMs still struggle with speaking and listening, teachers will need to blend AI-assisted writing with plenty of real conversation, peer review, and teacher feedback to ensure balanced language development.

Building Systems That Support Human Learning: at the institutional level, the challenge is about creating ecosystems where LLMs can be used ethically, inclusively, and sustainably. That starts with clear policy. Universities and educational institutions must define responsible AI use, including how to safeguard student information, deal with algorithmic prejudice, and uphold academic integrity when students utilize generative tools. Training is also important. For teachers to become proficient in using LLMs both technically and pedagogically, they require time, assistance, and professional development. This entails knowing how to assess AI results, incorporate it into teaching, and assist students in developing into critical, self-sufficient users.

Equity is another concern that can't be ignored. Many of the learners in this study were digitally confident and well-resourced—but that's not true everywhere. Institutions must ensure that LLM access doesn't become a new axis of inequality. That means investing in infrastructure, supporting multilingual learners, and designing tools that reflect a range of cultural and linguistic realities.

Finally, it’s important not to rush. The best way to introduce LLMs is gradually, through pilot programs, targeted training, and ongoing evaluation. This gives institutions a chance to learn from their learners, adjust their strategies, and make sure that AI enhances rather than disrupts the learning experience.People are just as important to this research's practical message as technology. LLMs can, in fact, speed up learning—but only if they are carefully presented, purposefully worded, and backed by sound human judgment. Tools that feel like teachers must be created by developers. Instead of teaching around AI, educators must teach with it. And for it to happen, institutions need to set up the infrastructure, policies, and space. LLMs can become more than just tools once these components are put together. They can develop into dependable, interesting, and empowering learning partners that support students in exploring, enjoying, and developing alongside language.

6.5 Limitations of the Study

Although this study offers valuable insights into how adult learners adopt and engage with Large Language Models in self-directed second language acquisition, several limitations must be acknowledged to appropriately contextualize its conclusions.

Cross-Sectional Design: The study used a cross-sectional survey design to record the attitudes, feelings, and actions of participants at one particular point in time. This method limits the ability to infer causality, even while it made data collection more efficient and helped spot early trends. For example, while the tutor framing and perceived usefulness of AI were substantially associated with sustained engagement, it is still unclear whether these beliefs influence or drive usage. AI may be more positively rated in hindsight by learners who are already motivated to utilize and trust it, which could indicate reverse causality or cognitive consistency effects. To investigate how these dynamics change over time and if the associations that have been observed remain true, more longitudinal research is required.

Reliance on Self-Report Measures: the study relied entirely on self-reported data for assessing key constructs such as trust, emotional engagement, and perceived learning gains. Although internal reliability was acceptable (e.g., Cronbach’s α around 0.75), self-reporting is inherently vulnerable to biases like social desirability, recall errors, and the halo effect. Subjective impressions of progress may not align with actual improvements in language proficiency. To enhance the robustness of findings, future work should integrate objective measures, such as standardized tests, system-generated usage logs, and third-party evaluations of learner output, including instructor-rated assignments.

Sampling Constraints and Generalizability: The participants were mostly self-selected adult learners from Anglophone and Central and Eastern European nations who were adept in digital technology. The findings' wider relevance is hampered by this relatively small demographic reach, especially for younger students, older adults, people with less access to digital technology, and people in underprivileged educational environments. Furthermore, the majority of participants utilized ChatGPT nearly exclusively, which limited the study's capacity to evaluate various LLM platforms. To get a more complete picture of how different learner groups experience LLM-supported language learning, broader, more diversified sampling will be necessary.

Short-Term Engagement and Novelty Effects: the study concentrated on users who were just beginning to engage with LLMs, a time when novelty was probably at its highest. It's possible that initial excitement rather than sustained involvement is the source of positive emotional reactions like enhanced motivation and enjoyment. It is unclear if these emotive improvements last or diminish as students get used to the technology in the absence of longitudinal follow-up. Future studies should look at how trust and motivation evolve over time and how different stages of AI integration into learning processes affect them.

In summary, although this study provides a valuable foundation for exploring the dynamics of AI use in language learning, its methodological limitations highlight the need for more thorough, extended, and varied research. Testing the TAM-Plus model's resilience and improving upcoming frameworks for AI-enhanced language instruction will depend on addressing these drawbacks.

6.6 Recommendations for Future Research

While this study offers insight into how adult learners interact with large language models in the context of second language acquisition, it also brings to light several critical avenues for further investigation. A deeper, more inclusive understanding of AI's pedagogical role will require research that moves beyond cross-sectional snapshots and self-reported experiences, focusing instead on how learner engagement evolves across time, populations, and instructional contexts. The requirement for longitudinal study is one strong argument. An early stage of learner interaction, marked by novelty, interest, and first emotional investment, is captured by the current findings. There is uncertainty over whether these early patterns continue and whether the seeming gains translate into sustained motivation and performance. In addition to regular self-assessments, future research should span multiple academic terms and include objective indicators like comprehensive usage records, standardized competence tests, and analyses of prompt construction quality. These long-term plans would demonstrate whether learners' initial enthusiasm fades in the absence of more comprehensive pedagogical integration or whether their talents grow as they continue to use AI. Equally important is the shift toward experimental methods that can clarify causality. The study could not separate the direction or influence of these variables, but it did find significant correlations between AI role framing, emotional involvement, and self-perceived learning. Randomized treatments, including putting students in alternative AI framings (peer versus tutor), or providing quick literacy and critical verification training, could provide tangible information on which teaching methods give the best results. Establishing causal relationships would give educators and developers the evidence they need to design pedagogically sound and empirically backed AI-enhanced environments. Another important study goal is to increase the demographic and geographic reach of participants. The current sample offers useful information, but it does not accurately reflect the variety of global language learners because it is mainly composed of young, technologically educated people from specific geographic areas. A greater range of users, including speakers of less widely taught languages, older students, children enrolled in formal school systems, and those with limited internet access, must be included in future research. To make sure that future tools work inclusive, flexible, across learning contexts, it will be important to take a look into how these learners view and engage with AI across various cultural, educational, and infrastructure frameworks.

Last but not least, creating AI tools that target oral language proficiency is still a major obstacle. This study revealed a significant gap in speaking and listening practice, which is probably related to the fact that current LLMs are predominantly text-based, even though it also emphasized reported gains in writing and grammar skills. Research should investigate whether advancements in speech recognition technology and voice-enabled models can effectively support spoken fluency, pronunciation, and real-time conversational interaction. Improving the multimodal capabilities of AI tools could help bridge the divide between written and spoken language practice, making the technology more comprehensive and communicatively relevant.

In summary, long-term participation, causal validation, increased participant diversity, and the creation of multimodal tools must be given top priority in future studies. By taking these actions, the field will be able to progress past early-stage insight and move closer to a more inclusive, transformative, and sustainable vision for AI-driven language instruction.

6.7 Concluding Remarks

This thesis set out to explore how adult language learners are engaging with Large Language Models (LLMs), like ChatGPT, in their self-directed pursuit of second language acquisition. The goal was to look beyond just whether these tools are easy to use, and instead ask: how do learners perceive them, how do they feel while using them, and what shapes their decision to keep coming back? What emerged is a more nuanced understanding of how technology and psychology intersect in educational settings captured in the proposed “TAM-Plus” framework, which expands on the traditional Technology Acceptance Model by adding three key dimensions: how learners conceptualize the AI’s role, how much they trust it (critically), and how emotionally engaged they feel when using it. The findings make one thing clear: the value of LLMs as learning tools doesn’t come just from their technical capabilities. Instead, their effectiveness is shaped by the learner’s mindset by whether they see the AI as a helpful tutor, feel motivated and confident while using it, and remain thoughtful enough to double-check its suggestions. In particular, framing the AI as a tutor not just a tool or chat partner was shown to increase trust and boost perceived learning, mostly because it made the learning experience more emotionally positive. This highlights how much social perception matters when we introduce AI into education. Equally important was the finding that learners don’t blindly follow what the AI tells them. Most took a thoughtful, skeptical approach what we’ve called “critical trust.” They valued the help, but kept their own judgment in the loop. This attitude points to a new kind of literacy that goes hand-in-hand with language learning today: understanding how to interact with, evaluate, and make the most of AI systems.

Taken together, this research encourages us to view LLMs not just as smart tools, but as socially meaningful actors in the learning environment. Their value isn't automatic—it depends on how learners relate to them, what kind of support they offer, and how those interactions are emotionally and cognitively experienced. Designing for these human factors is key. With the right framing, support, and safeguards, LLMs can become powerful allies in language learning: offering personalized practice, reducing anxiety, and creating space for deeper reflection and creativity. Looking ahead, the real opportunity is to embrace these technologies without losing sight of what matters most in education: the human connection, the joy of discovery, and the meaningful use of language. If we keep those priorities in focus, LLMs can help amplify, not replace, the best parts of what it means to learn a language.

References

1. AI-enabled language speaking coaching for dual language learners. (2019). IADIS International Journal on WWW/Internet, 17(1).https://doi.org/10.33965/ijwi_2019171105
2. Akhiat, M. (2024). Second Language Acquisition in the Era of Technology and Artificial Intelligence: Exploring New Frontiers.
3. Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using Chatbots as AI Conversational Partners in Language Learning. *Applied Sciences*, *12*(17), 8427. https://doi.org/10.3390/app12178427
4. Camilleri, M. A. (2024). Factors affecting performance expectancy and intentions to use ChatGPT: Using SmartPLS to advance an information technology acceptance framework. *Technological Forecasting and Social Change*, *201*, 123247. https://doi.org/10.1016/j.techfore.2024.123247
5. Dam, S. K., Hong, C. S., Qiao, Y., & Zhang, C. (2024). *A Complete Survey on LLM-based AI Chatbots* (No. arXiv:2406.16937). arXiv. https://doi.org/10.48550/arXiv.2406.16937
6. Graesser, A. C., Li, H., & Forsyth, C. (2014). Learning by Communicating in Natural Language With Conversational Agents. *Current Directions in Psychological Science*, *23*(5), 374–380. https://doi.org/10.1177/0963721414540680
7. Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, *55*(1), 48–63. https://doi.org/10.1080/15391523.2022.2142873
8. Jian, M. J. K. O. (2023). Personalized learning through AI. *Advances in Engineering Innovation*, *5*(1), 16–19. https://doi.org/10.54254/2977-3903/5/2023039
9. Kim, W.-H., & Kim, J.-H. (2020). Individualized AI Tutor Based on Developmental Learning Networks. *IEEE Access*, *8*, 27927–27937. https://doi.org/10.1109/ACCESS.2020.2972167
10. Krashen, S. D. (1984). Principles and practice in second language acquisition (Reprinted). Pergamon Press.

11. Lemon, O. (2022). Conversational AI for multi-agent communication in Natural Language: Research directions at the Interaction Lab. *AI Communications*, *35*(4), 295–308. https://doi.org/10.3233/AIC-220147
12. Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI Chatbots for Content and Language Integrated Learning. *Applied Sciences*, *12*(7), 3239. https://doi.org/10.3390/app12073239
13. Ngo, K. T. (2024). The Use of ChatGPT for Vocabulary Acquisition: A Literature Review. *International Journal of AI in Language Education*, *1*(2), 1–17. https://doi.org/10.54855/ijaile.24121
14. Pan, M., Kitson, A., Wan, H., & Prpa, M. (2024). ELLMA-T: An Embodied LLM-agent for Supporting English Language Learning in Social VR (No. arXiv:2410.02406). arXiv. https://doi.org/10.48550/arXiv.2410.02406
15. Park, H.-Y. (2023). Application of ChatGPT for an English Learning Platform. *STEM Journal*, *24*(3), 30–48. https://doi.org/10.16875/stem.2023.24.3.30
16. Pratiwi, N., Efendy, A. G., Rini, H. C., & Ahmed, N. A. (2024). Speaking Practice using ChatGPT's Voice Conversation: A Review on Potentials and Concerns. *Journal of Language Intelligence and Culture*, *6*(1), 59–72. https://doi.org/10.35719/jlic.v6i1.149
17. Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, *14*, 1255594. https://doi.org/10.3389/fpsyg.2023.1255594
18. Settles, B., Brust, C., Gustafson, E., Hagiwara, M., & Madnani, N. (2018). Second Language Acquisition Modeling. *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications*, 56–65. https://doi.org/10.18653/v1/W18-0506
19. Tarone, E. (1974). SPEECH PERCEPTION IN SECOND LANGUAGE ACQUISITION: A SUGGESTED MODEL. *Language Learning*, *24*(2), 223–233. https://doi.org/10.1111/j.1467-1770.1974.tb00504.x
20. Theophilou, E., Koyutürk, C., Yavari, M., Bursic, S., Donabauer, G., Telari, A., Testa, A., Boiano, R., Hernandez-Leo, D., Ruskov, M., Taibi, D., Gabbiadini, A., & Ognibene, D. (2023). Learning to Prompt in the Classroom to Understand AI Limits: A

Pilot Study. In R. Basili, D. Lembo, C. Limongelli, & A. Orlandini (Eds.), *AIxIA 2023 – Advances in Artificial Intelligence* (Vol. 14318, pp. 481–496). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47546-7_33

21. Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01000
22. Wei, Z. (n.d.). Chat GPT Integrated with Voice Assistant as Learning Oral Chat-based Constructive Communication to Improve Communicative Competence for EFL Learners.
23. Willie Moore & Li-Shiang Tsay. (2024). From Data to Decisions: Leveraging AI for Proactive Education Strategies. *International Conference on AI Research*, *4*(1), 281–288. https://doi.org/10.34190/icair.4.1.3082
24. Wu, M. M. (2010). An Exploratory Study of the Language-learning Style Preferences and Language-learning Motivation of ESL Learners at a Vocational Education Institute in Hong Kong.
25. Xiao, Y., & Zhi, Y. (2023). An Exploratory Study of EFL Learners' Use of ChatGPT for Language Learning Tasks: Experience and Perceptions. *Languages*, *8*(3), 212. https://doi.org/10.3390/languages8030212
26. Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, *10*(3), e25370. https://doi.org/10.1016/j.heliyon.2024.e25370
27. Ziaja, R. (2024). Korzyści i granice wykorzystania sztucznej inteligencji na lekcjach języka niemieckiego na przykładzie ChatGPT. *Neofilolog*, *62/2*, 521–540. https://doi.org/10.14746/n.2024.62.2.11

Appendix

Appendix A: Survey

QR Code:

Survey:

Learner Perceptions & Engagement Survey

You are invited to participate in a research study exploring how learners use AI tools like ChatGPT for language learning. This survey will take **10 minutes**.

Your responses are **anonymous** and **voluntary**. By proceeding, you confirm you are **18+** and agree to participate.

Consent

* I have read and agree to participate.

Background Information

* Have you used an AI language model (e.g., ChatGPT, Duolingo Max, Gemini) for practicing a second language?

- Yes
- No

Please provide basic information about yourself.

* Age

- 18–24
- 25–34
- 35–44
- 45+

* Gender

- Male
- Female
- Non-binary/Other
- Prefer not to say

* First/Native Language:

* Target Language(s) You Are Learning:

* Self-Rated Proficiency Level:

- Beginner (A1)
- Elementary (A2)
- Intermediate (B1)
- Upper-Intermediate (B2)
- Advanced (C1/C2)

AI Tool Usage

* Which AI tools have you used for language learning? *(Check all that apply)*

- ChatGPT
- Gemini
- DeepSeek
- Duolingo Max (AI features)

- Other: _______________

* How long have you used AI tools for language learning?

- <1 month
- 1–3 months
- 4–6 months
- 6 months+

* How often do you use AI tools specifically to support your language learning?

- Every time
- Every second session
- Every third session
- Rarely

* What activities do you use AI tools for? *(Check all that apply)*

- Writing practice (e.g., essays, sentences)
- Speaking practice
- Grammar explanations
- Vocabulary learning (e.g., understanding word meanings, usage in context)
- Translation
- Pronunciation/listening (if voice-enabled)
- Other: _______________

* Average AI session length:

- <10 minutes
- 10–20 minutes
- 20–30 minutes
- 30 minutes +

Perceptions of AI Use

Rate your agreement with the following statements *(1 = Strongly Disagree, 5 = Strongly Agree)*

Statement	1	2	3	4	5
I view the AI as a tutor that teaches me new things.					
I see the AI as a peer to practice with.					
The AI is just a tool I use when I need					
I trust the AI's corrections					
I sometimes doubt the AI's accuracy and verify its answers.					
Using the AI improves my language					
The AI is easy to use and navigate.					
I plan to continue using AI tools for language learning.					
People around me (teachers/ friends) encourage AI					
I feel less anxious practicing with the AI than with					
Using the AI has increased my confidence in the language.					

Statement	1	2	3	4	5
I find interacting with the AI enjoyable.					
The AI motivates me to practice more frequently.					

Skill Development

How has the AI impacted your skills? *(1 = No Improvement, 5 = Significant Improvement)*

Skill	1	2	3	4	5
Vocabulary					
Grammar					
Writing fluency					
Speaking confidence					
Listening comprehensio					

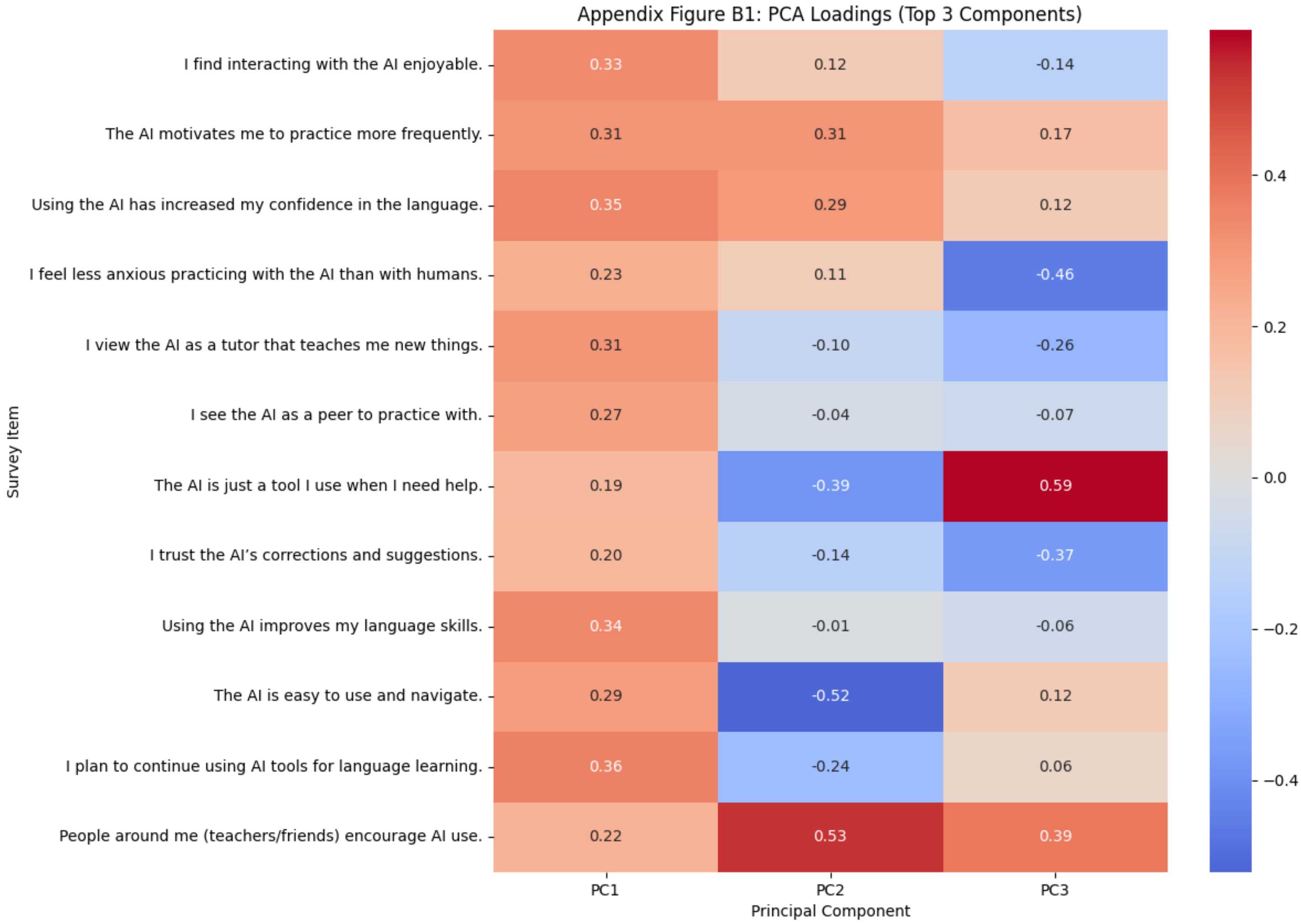
Open Questions

1. What do you like most about using AI for language learning?

2. What challenges have you faced when using AI?

Thank you for participating! Your responses will help improve AI tools for language learners.

Appendix B: PCA Loadings (Top 3 Components)



Appendix C: PCA Loadings (First 3 Components)