



ABSTRACT
SUBMISSION
DEADLINE

30 SEPTEMBER



AELTE'25

İNÖED FOURTH INTERNATIONAL CONFERENCE

10-11-OCTOBER 2025

"ARTIFICIAL INTELLIGENCE PRACTICES IN FOREIGN LANGUAGE EDUCATION"



John Silver
U.S Embassy



Vincent Greenier
Keynote Speech
Ethical Use of Generative
AI in Research Literacy
in EFL
American University of Sharjah (AUS)



Assoc. Prof. Dr. Hamed
Barjesteh
Islamic Azad University, Ayatollah
Ameli Branch



Prof. Dr. Radka Hrdz
Keynote Speech
Ethical AI in Language Education:
Contesting Cars,
Culture, and Human Connection

SOUTHERN CONNECTICUT
STATE UNIVERSITY



Dr. Nidya Tsimis
Keynote Speech
Designing for Change:
Adapting the EFL Curricula
for a More Sustainable World

UNIVERSITY OF TURKISH
AERONAUTICAL ASSOCIATION



Prof. Dr. Ahmet İbrahim ÇAKIR
Keynote Speech
The ways in which
AI is changing our EFL practices

UFUK UNIVERSITY



Dr. Esra MİTİR
Keynote Speech
ID Thinking: Cultivating a Critical
Mindset in EFL

ÇANKAYA UNIVERSITY



Assoc. Prof. Dr. Sennur İbrahim
Keynote Speech
Investigating Teaching
Turkish as a Heritage
Language in Germany
From the Perspective of English Teachers

AUGSBURG UNIVERSITY



CONFERENCE PROCEEDINGS

This publication is the Book of Proceedings of the international AELTE 4th Conference, held on 10–11 October 2025 at Kahramanmaraş Sütçü İmam University, Kahramanmaraş, Türkiye, in collaboration with AELTE and hosted by Kahramanmaraş Sütçü İmam University. All legal and academic responsibilities regarding the articles belong to the authors.

DECEMBER 2025

ISBN-978-605-73434-7-5

EDITORS

Prof. Dr. Arif SARIÇOBAN

Dr. Öğr. Üyesi Aylin YARDIMCI

Dr. Öğr. Üyesi Mustafa ÇAKIR



KAHRAMANMARAŞ
SÜTÇÜ İMAM UNIVERSITY



ASSOCIATION OF
ENGLISH LANGUAGE
TEACHER EDUCATORS



KAHRAMANMARAŞ
SÜTÇÜ İMAM UNIVERSITY
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
DEPARTMENT OF TRANSLATION AND INTERPRETATION

İNÖED FOURTH INTERNATIONAL CONFERENCE

10-11-OCTOBER 2025

“ARTIFICIAL INTELLIGENCE PRACTICES IN FOREIGN LANGUAGE EDUCATION”

❖ Committees

President

Prof.Dr. Alptekin YASIM - Rector, Kahramanmaraş Sütçü İmam University

Partners

Prof. Dr. Arif Sarıçoban – President & Association of English Language Teacher Educators

Prof. Dr. Yakup POYRAZ - Dean, Kahramanmaraş Sütçü İmam University, Faculty of Humanities and Social Sciences

Organization Committee

Prof. Dr. Arif Sarıçoban - President, INOED & Senior Lecturer, Selçuk University, Turkey

Prof. Dr. Yakup Poyraz Dean, Faculty of Humanities and Social Sciences, Kahramanmaraş Sütçü İmam University, Türkiye

Doç. Dr. Ceyhun Karabıyık - Ufuk University, Turkey

Doç. Dr. Aydan Irgatoğlu - Ankara Hacı Bayram Veli University, Turkey

Dr. Öğr. Üyesi Aylin Yardımcı - Kahramanmaraş Sütçü İmam University, Turkey

Dr. Öğr. Üyesi Semih Sarıgül - Kahramanmaraş Sütçü İmam University, Turkey

Dr. Öğr. Üyesi Zeynep Canlı - Kahramanmaraş Sütçü İmam University, Turkey

Dr. Öğr. Üyesi Betül - Kahramanmaraş Sütçü İmam University, Turkey

Dr. Öğr. Üyesi Filiz Yörük Çevik - Kahramanmaraş Sütçü İmam University, Turkey

Dr. Öğr. Üyesi Hatice Ergül - Hacettepe University, Ankara

Dr. Esen Metin - Çankaya University, Ankara

Dr. Öğr. Üyesi Mustafa Çakır - Kahramanmaraş Sütçü İmam University, Turkey

National Scientific Board

Abdülvahit Çakır – Ufuk University, Turkey

Ahmet Çekiç – Cumhuriyet University, Turkey

Arif Bakla – Yıldırım Beyazıt University, Turkey

Aslı Özlem Tarakçıoğlu – Ankara Hacı Bayram Veli University, Turkey

Asuman Aşık – Gazi University, Turkey

Aydan Irgatoğlu – Ankara Hacı Bayram Veli University, Turkey

Aylin Yardımcı - Kahramanmaraş Sütçü İmam University, Turkey

Aysu Erden – Maltepe University, Turkey

Ayşegül Takkaç – Atatürk University, Turkey

Bahadır Cahit Tosun – Selçuk University, Turkey

Belgin Aydın – TED University, Turkey

Burak Telli - Kahramanmaraş Sütçü İmam University, Turkey

Bülent Sayak - Kahramanmaraş Sütçü İmam University, Turkey

Ceyhun Karabıyık – Ufuk University, Turkey

Ceyhun Yükselir – Osmaniye Korkut Ata University, Turkey

Cemal Çakır – Gazi University, Turkey

Ceylan Yangın Ersanlı – Ondokuz Mayıs University, Turkey

Dilek İnal – İstanbul University-Çerrahpaşa, Turkey

Dinçay Köksal – Çanakkale Onsekiz Mart University, Turkey

Doğan Saltaş – Ardahan University, Turkey

Elif Tokdemir Demirel – Kırıkkale University, Turkey

Emrah Cinkara – Gaziantep University, Turkey

Emrah Dolgunsöz – Bayburt University, Turkey

Erdem Akbaş – Erciyes University, Turkey

Erdoğan Bada – Hakkâri University, Turkey

Fatma Kalpaklı – Selçuk University, Turkey

İsmail Fırat Altay – Hacettepe University, Turkey

İskender Sarıgöz – Gazi University, Turkey

Gencer Elkılıç – Kafkas University, Turkey

Gonca Yangın Ekşi – Gazi University, Turkey

Gülsev Pakkan – Selçuk University, Turkey

Gürkan Dağbaşı – Ankara Hacı Bayram Veli University, Turkey
Güven Mengü – Ankara Hacı Bayram Veli University, Turkey
Hacer Hande Uysal – Hacettepe University, Turkey
Hakan Demiröz – Ankara Social Sciences University, Turkey
Hasan Çağlar Başol – Selçuk University, Turkey
Hasan Bedir – Çukurova University, Turkey
İrfan Tosuncuoglu – Karabük University, Turkey
Iryna Semeniuk Zümrütdal – Piri Reis University, Turkey
İsmail Çakır – Ankara Social Sciences University, Turkey
İsmail Hakkı Mirici – Hacettepe University, Turkey
Kadriye Dilek Bacanak – Gazi University, Turkey
Korkut Uluç İşisağ – Ankara Hacı Bayram Veli University, Turkey
Kürşat Cesur – Çanakkale Onsekiz Mart University, Turkey
Mehmet Bardakçı – Gaziantep University, Turkey
Mehmet Kılıç – Gaziantep University, Turkey
Mehmet Takkaç – Atatürk University, Turkey
Murat Özcan – Gazi University, Turkey
Murat Hişmanoğlu – Uşak University, Turkey
Mustafa Çakır – Anadolu University, Turkey
Mustafa Naci Kayaoglu – Karadeniz Teknik University, Turkey
Mustafa Zeki Çıraklı – Karadeniz Teknik University, Turkey
Nalan Kızıltan – Ondokuz Mayıs University, Turkey
Nazlı Gündüz – Ankara Hacı Bayram Veli University, Turkey
Nejla Gezmiş – Kırıkkale University, Turkey
Oktay Akarsu – Atatürk University, Turkey
Oktay Yağız – Atatürk University, Turkey
Oya Tunaboylu – Süleyman Demirel University, Turkey
Sadi Gedik - Kahramanmaraş Sütçü İmam University, Turkey
Saliha Defne Erdem Mete – Selçuk University, Turkey
Selim Somuncu - Kahramanmaraş Sütçü İmam University, Turkey
Selma Elyıldırım – Ankara Hacı Bayram Veli University, Turkey
Selmin Söylemez – Ankara Hacı Bayram Veli University, Turkey
Yasin Murat Demir – Ankara Hacı Bayram Veli University, Turkey
Muammer Sarıkaya – Ankara Hacı Bayram Veli University, Turkey
Yakup Poyraz - Kahramanmaraş Sütçü İmam University, Turkey
Yiğit Sümbül – Ankara Hacı Bayram Veli University, Turkey
Zekiye Müge Tavil – Gazi University, Turkey
Zennure Ergün Gündüz – Ardahan University, Turkey
Zeynep Arkan – Ankara Hacı Bayram Veli University, Turkey
Parisa Çoşkun – Karatay University, Turkey

International Scientific Board

Ahmar Mahboob – The University of Sydney, Australia
Ambreen Safdar Kharbe – Najran University, Saudi Arabia
Anette Ipsen – University College Copenhagen, Denmark
Badriyah El-Daw – Lebanese University, Lebanon
Christopher Cairney – Middle Georgia State University, USA
Iryna Semeniuk Zümrütdal – Piri Reis University
Jerome Bush – USA
Jesus Garcia-Laborda – Alcala University, Spain
Jerome Bush – USA
Lotte Lindberg – University College Copenhagen, Denmark
Maggie Sokolik – University of California, Berkeley, United States of America
Manana Rusieshvili-Cartledge – Tbilisi State University, Georgia

María Jesús Sánchez – University of Salamanca, Spain

Miriam Eisenstein – New York University, United States of America

Oksana Chaika – National University of Life And Environmental Sciences of Ukraine, Ukraine

Rome Aboh – University of Uyo, Nigeria

Salam Yusuf Nuhu Inuwa – Kano State College Of Arts And Sciences, Nigeria

Sejdi M. Gashi – Institute of Albanology-Pristina (Kosovo), Albania

Stefan Rathert - Kahramanmaraş Sütçü İmam University, Turkey

Zeleke Arficho Ayele – Hawassa University, Ethiopia



KAHRAMANMARAŞ
SÜTÇÜ İMAM UNIVERSITY



ASSOCIATION OF
ENGLISH LANGUAGE
TEACHER EDUCATORS



KAHRAMANMARAŞ
SÜTÇÜ İMAM UNIVERSITY
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
DEPARTMENT OF TRANSLATION AND INTERPRETATION

İNÖED FOURTH INTERNATIONAL CONFERENCE

10-11-OCTOBER 2025

“ARTIFICAL INTELLIGENCE PRACTICES IN FOREIGN LANGUAGE EDUCATION”

❖ Table of Contents

Author	Title	Page
Inamul AZAD Azra TAJHIZI	Code-Switching with AI: Teaching Multilingual Minds in the Digital Age	9
Nazila MOTAHARI Sonia BATRA	Digital Applications in Foreign Language Education	15
Aydan IRGATOĞLU	Boosting EFL Learners' Motivation through Online Project-Based Learning	18
Aydan IRGATOĞLU	How Digital Literacies Shape EFL Learning: Reducing Stress, Easing Anxiety, and Boosting Enjoyment	28
Yunus Emre AVCI Ayhan YAVUZ	English Teaching Department Teacher Candidates' Employment Perceptions And Public–Private Sector Preferences	35
Selma ELYILDIRIM	Feedback Provision: Which one is Better, AI or Human Grading?	43
Selma ELYILDIRIM	Artificial Intelligence in the Translation of Idiomatic Expressions: To What Extent it is Reliable?	54
Mehmet TUNAZ	School Climate as a Moderator in the Relationship Between Job Satisfaction, Teacher Immunity, and Burnout: A Proposed Moderated Mediation Model	64
Fatma ÇILOĞLAN KONUR	Conjunctive Devices in Undergraduate Literature Reviews: A Corpus-Based Study	70
Zeynep DAŞER Bahar KAR	EFL learners' frequency and perceptions of using ChatGPT in informal digital learning of English: Does personality matter?	78
Aysun Barış KOSAOĞULLARI Vasfiye GEÇKİN	From Red Pen to Algorithm: A Systematic Review of AI and Teacher Feedback in Writing	99
Güven MENGÜ	Teaching Prosody through Integrated Skills: An AI-Supported Pedagogical Perspective	117
Merve ELBİRLİK TÜLEK	Integrating AI into Undergraduate Research: Exploring Students' Perceptions, Practices, and Ethics	130
Hossein ISAAE Hamehnaz GHOLIPOUR Hamed BARJESTEH	AI and Multisensory Language Learning: Unlocking Richer Pathways in L2 Acquisition	138
Şahin GÖK	The Role of Artificial Intelligence Tools in Enhancing English Proficiency and Test Preparation: Evidence from Freshmen Literature Students at İstanbul Gelişim University	153
Nazlı GÜNDÜZ	Exploring EFL Prep Students' Attitudes Towards Artificial Intelligence and Student Willingness to Communicate in Language Learning	159
Ziyoda KHALMATOVA	Exploring EFL Students' Attitudes Toward Artificial Intelligence Tools in Language Learning: A Qualitative Exploration	168
Abdülvahit ÇAKIR	The Ways in which AI is Changing Our ELT Practices	179
Ahmet Selçuk AKDEMİR	A Review of Willingness to Communicate (WTC) in L2 and Learner Psychology	188
Ahmet Selçuk AKDEMİR	Cultural Elements of 'Uptown': A Coursebook Analysis	192

Code switching with AI: Teaching Multilingual Minds in the digital Age

Dr. Inamul Azad

Central University of Karnataka, Kalaburagi, India

Abstract

In the rapidly evolving landscape of digital education, code-switching, the fluid shifting between two or more languages, has gained renewed significance through the integration of Artificial Intelligence (AI) in multilingual classrooms. This paper, "Code Switching with AI: Teaching Multilingual Minds in the Digital Age," explores how AI-enabled tools transform linguistic flexibility into a pedagogical resource rather than a barrier. Drawing from sociolinguistics, applied linguistics, and AI-mediated learning theories, the study argues that AI can amplify the cognitive, cultural, and communicative benefits of code-switching when used intentionally by educators. AI applications, such as adaptive chatbots, multilingual feedback systems, voice-recognition platforms, and translation-supported learning environments, facilitate smoother transitions between languages, reduce learner anxiety, and promote deeper comprehension for students navigating multiple linguistic identities.

The paper highlights how AI supports dynamic bilingualism by personalizing instruction, recognizing learners' language patterns, and providing scaffolded explanations across languages in real time. It also examines potential risks, including over-dependence on machine translation, cultural flattening, and the marginalization of minority languages, emphasizing the need for pedagogical frameworks that prioritize critical digital literacy and equitable language representation. Methodologically, the study adopts a mixed-methods approach, combining classroom observations, AI-assisted task analyses, and interviews with multilingual learners and teachers.

Findings suggest that when guided by informed instructional choices, AI-supported code-switching enhances cognitive processing, fosters inclusive participation, and strengthens language retention in multilingual educational spaces. The paper concludes by proposing an AI-responsive multilingual pedagogy that positions code-switching as a legitimate, creative, and empowering strategy for teaching multilingual minds in the digital age.

Keywords: Code-switching; Artificial Intelligence (AI); Multilingual Education; Digital Pedagogy; Language Learning

Introduction

The digital age has transformed the way languages are learned, taught, and used in everyday communication. As classrooms grow increasingly multilingual, students frequently draw on more than one language to make sense of new ideas, express complex thoughts, or bridge gaps in understanding. This linguistic flexibility, commonly known as code-switching, has traditionally been viewed with suspicion in formal education, where monolingual norms dominated classroom discourse. However, the rapid integration of Artificial Intelligence (AI) into learning environments has disrupted these assumptions. AI-powered tools, from automated translators to adaptive learning systems, now interact with students in multiple languages, often mirroring the natural rhythm of multilingual speech.

The growing presence of AI raises important pedagogical questions: Can AI make code-switching a more purposeful instructional strategy? How do multilingual learners engage with AI platforms that respond to them in different languages? Moreover, most crucially, what does it mean to teach multilingual minds in a world where human and machine languages intersect? This paper explores these questions by examining how AI reshapes the use, function, and cultural meaning of code-switching in contemporary classrooms. By foregrounding student experiences and teacher perspectives, the study aims to understand whether AI can help educators move from viewing code-switching as an obstacle to recognizing it as a dynamic learning resource.

Literature Review

Research on code-switching has long emphasized its cognitive and communicative value. Early sociolinguistic studies (Gumperz, 1982; Myers-Scotton, 1993) identified code-switching as a deliberate and meaningful practice rather than a sign of linguistic deficiency. In multilingual classrooms, scholars such as García (2009) and Creese & Blackledge (2010) introduced the idea of translanguaging, arguing that multilingual learners draw on their full linguistic repertoires to make meaning. These studies collectively reframed code-switching as a natural strategy that supports comprehension, expression of identity, and classroom participation.

Parallel to this, research on digital learning and AI-mediated education has expanded rapidly. AI applications, such as language models, translation software, and speech-recognition systems, have been shown to personalize instruction and offer real-time linguistic support (Luckin et al., 2016; Holmes et al., 2019). Recent work also suggests that AI tools can help scaffold multilingual interactions, particularly for students navigating hybrid linguistic spaces (Zou & Xie, 2021). Studies on machine translation in language classrooms (Lee, 2022) demonstrate both its potential to support more profound understanding and its risks, including learner over-reliance and the loss of linguistic nuance.

Despite these developments, there is a limited amount of research examining the intersection of AI and code-switching. Most discussions treat the two areas separately, leaving a gap concerning how AI might actively enable or transform code-switching practices. This emerging intersection warrants attention, particularly as educational institutions rapidly adopt AI tools. The present study situates itself at this crossroads, exploring how AI technologies influence the ways multilingual learners shift between languages and how teachers make pedagogical sense of these shifts.

Research Questions

To address the gaps identified in the literature, the study is guided by the following research questions:

- How do multilingual learners engage in code-switching when interacting with AI-powered educational tools?
- In what ways do AI applications support, shape, or constrain code-switching as a learning strategy in multilingual classrooms?
- How do teachers perceive the role of AI in facilitating or disrupting code-switching practices during instruction?
- What pedagogical approaches can help integrate AI-supported code-switching into meaningful, equitable multilingual teaching?

2. Article Structure

This article is organised into six major sections to provide a coherent and comprehensive examination of AI-supported code-switching in multilingual education. The introduction outlines the conceptual background and significance of the study. The literature review synthesizes existing scholarship on code-switching, translanguaging, and AI-based learning, highlighting the need to study the convergence of these domains. The methodology section explains the research design, participant profile, data sources, and analytical procedures. The results section presents empirical patterns observed across classrooms, AI interactions, and participant interviews. The discussion interprets these findings in relation to existing theories of multilingual learning and human–machine interaction. Finally, the conclusion summarises the main contributions, identifies potential risks, and proposes future directions for AI-responsive multilingual pedagogy.

3. Method

This study followed a mixed-methods exploratory design, combining qualitative and quantitative techniques to capture the complexity of multilingual learners' interactions with AI tools.

3.1 Participants and Setting

The research was conducted across three urban schools and one university language centre where English, Hindi, Urdu, and regional languages commonly coexisted. A total of 92 multilingual learners aged 14-22 and 11 language teachers participated. All institutions had recently integrated AI-based learning platforms, including multilingual chatbots, automated translation interfaces, and speech-recognition applications.

3.2 Data Sources

Data were collected from three primary sources:

Classroom Observations

Twenty-one sessions were observed over an eight-week period. These sessions documented natural instances of code-switching during AI-assisted tasks.

AI Interaction Logs

Exported logs from AI tools captured real-time language switching patterns, translation usage, and system-led prompts. These logs provided quantitative indicators such as frequency and distribution of switches.

Semi-Structured Interviews

In-depth interviews were conducted with 36 students and all 11 teachers. Questions explored learner attitudes toward AI, comfort with switching languages, and teachers' perceptions of AI's pedagogical value.

3.3 Data Analysis

Qualitative data (observations and interviews) were coded thematically following Braun & Clarke's (2006) reflexive approach. Codes included "clarification through switching," "identity-based switching," "AI-triggered switching," and "loss of nuance in translation."

Quantitative patterns from interaction logs were analyzed using descriptive statistics to identify the frequency of switches, the average length of bilingual exchanges, and the comparative use of L1 and L2 across tasks.

Ethical approval was obtained, and all participants provided informed consent.

4. Results

Four significant empirical findings emerged from the analysis:

4.1 AI Tools Increased the Frequency of Purposeful Code-Switching

Across all sites, AI interfaces prompted more natural and confident switching. Students frequently moved between English and their home languages to clarify meanings or generate ideas. AI translation and paraphrasing tools appeared to reduce anxiety around “making mistakes,” a theme widely noted in the interviews. This aligns with existing evidence that digital scaffolds reduce cognitive load in multilingual learning (Zou & Xie, 2021).

4.2 Students Used Code-Switching to Negotiate Meaning with AI

Interaction logs revealed that 61% of multilingual learners intentionally alternated languages to check comprehension. For example, students typed a sentence in English and then repeated the meaning in Hindi or Urdu to see how the AI responded. This double-checking behaviour reflects what García (2009) identifies as translanguaging for sense-making.

4.3 Teachers Reported Both Benefits and Pedagogical Concerns

Teachers appreciated AI’s ability to provide instant explanations, pronunciation support, and multi-language feedback. However, many expressed concerns about over-reliance on machine translation, echoing earlier warnings by Lee (2022) that excessive use of digital translation may reduce productive language skills. Teachers also worried that AI sometimes flattened cultural nuances or provided overly literal translations.

4.4 Minority Languages Benefited Less from AI Support

Students using languages with limited AI training data, such as regional dialects, showed lower accuracy and lagging comprehension support. This confirms patterns observed by Joshi et al. (2020), who argue that low-resource languages remain underrepresented in AI systems.

5. Discussion

This study demonstrates that AI does not merely assist multilingual learning, but it actively reshapes how learners switch between languages. The findings align with sociocultural perspectives that view language use as a fluid and adaptive resource (Creese & Blackledge, 2010). AI-enabled tools amplify this fluidity by offering real-time scaffolding, ultimately normalizing code-switching in educational spaces.

5.1 Reframing Code-Switching as a Cognitive Strategy

Students used AI to validate their bilingual expressions, reflecting a cognitive self-monitoring process similar to what Myers-Scotton (1993) described as “markedness evaluation.” AI interacted with these switches, reinforcing that multiple languages can coexist productively in learning tasks. This contradicts traditional monolingual classroom norms and supports emerging research that recognizes bilingual flexibility as a cognitive asset (Bialystok, 2018).

5.2 AI as a “Translanguaging Partner.”

The notion of AI acting as a translanguaging partner emerged repeatedly. Students treated AI systems not merely as tools but as interactive agents capable of participating in bilingual conversations. This highlights a new field intersection: human machine code-switching. It suggests the need for pedagogies that position AI as part of multilingual meaning-making rather than a replacement for human instruction.

5.3 Pedagogical Risks

Despite its benefits, AI introduces significant challenges. Over-reliance on translation may discourage productive language use, and inconsistent support for minority languages may further widen linguistic inequalities. Additionally, AI's occasional inability to interpret pragmatic meanings raises concerns about cultural flattening, a theme consistent with research on algorithmic bias in multilingual AI (Blodgett et al., 2020).

5.4 Implications for Teachers

Teachers need structured frameworks for integrating AI into multilingual instruction. Training should help educators balance AI assistance with classical language teaching methods so that code-switching remains purposeful, not compensatory. The study suggests adopting a critical digital multilingualism approach, where students evaluate AI translations and discuss inaccuracies to deepen linguistic awareness.

6. Conclusions

This study contributes to the growing scholarship on multilingual education by demonstrating that AI meaningfully transforms the function of code-switching in classroom settings. When used intentionally, AI tools encourage linguistic experimentation, support comprehension, and reduce anxiety, especially among learners who navigate complex bilingual or multilingual identities. However, the findings also caution against the uncritical adoption of this approach. AI may unintentionally promote linguistic inequity, flatten cultural nuance, or trigger over-dependence on machine-assisted language use.

To move forward, educators must adopt pedagogical models that position AI as a supportive partner rather than as an authority. Future research should investigate the long-term effects of AI-supported code-switching, the differences between low-resource and high-resource language users, and the ethical implications of AI-mediated linguistic practices. Ultimately, teaching multilingual minds in the digital age requires aligning technology with human-centered, culturally rooted educational goals.

References

Bialystok, E. (2018). Bilingual education for young children: Review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism*, 21(6), 666-679.

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of “bias” in natural language processing. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5454-5476.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Creese, A., & Blackledge, A. (2010). Translanguaging in the bilingual classroom: A pedagogy for learning and teaching? *The Modern Language Journal*, 94(1), 103-115.

García, O. (2009). Bilingual education in the 21st century: A global perspective. Wiley-Blackwell.

Gumperz, J. J. (1982). Discourse strategies. Cambridge University Press.

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6282–6293.

Lee, J. (2022). Machine translation in language education: Opportunities and challenges. *Language Learning & Technology*, 26(1), 1–23.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.

Myers-Scotton, C. (1993). Social motivations for codeswitching: Evidence from Africa. Oxford University Press.

Zou, D., & Xie, H. (2021). Digital scaffolding for supporting language learning in multilingual contexts: A review. *Computer Assisted Language Learning*, 34(7), 841-872.

Digital Applications in Language Learning Education: Enhancing ESL Interview Preparation through a Customized GPT

Nazila Motahari & Sonia Batra

Higher Colleges of Technology, Abu Dhabi, UAE

APA Citation:

Motahari N. & Batra S. (2025).Digital Applications in Language Learning Education: Enhancing ESL Interview Preparation through a Customized GPT

Abstract

As artificial intelligence (AI) tools rapidly transform education, language instructors are exploring new ways to integrate digital applications into teaching and learning.

This paper presents a practical case study on the use of a customized GPT-based AI tool to support ESL students in a Professional Spoken Communication course. The customized AI was designed to simulate interview scenarios, provide real-time feedback, and guide students through structured interview practice sessions. The project aimed to enhance students' fluency, confidence, and readiness for professional communication assessments.

Findings highlight significant improvements in learner engagement and self-efficacy, alongside valuable pedagogical insights for educators seeking to blend AI-driven tools with traditional instruction. The paper concludes by reflecting on lessons learned and outlining future directions for AI-enhanced language learning.

Introduction

Digital transformation has reshaped nearly every aspect of education, and language teaching is no exception. As ESL educators, we continually seek new strategies to engage students, personalize instruction, and extend learning beyond the classroom.

The rise of artificial intelligence, particularly in natural language processing, offers powerful tools for achieving these goals. Among these, customized GPT (Generative Pre-trained Transformer) models within ChatGPT provide opportunities to create interactive, responsive, and context-specific learning applications.

This paper reports on the development and classroom integration of a customized GPT tool designed to help students in the Professional Spoken Communication course prepare for their interview examinations.

The initiative emerged from a pedagogical challenge familiar to many language instructors: how to provide ample speaking practice and individualized feedback within time-limited class sessions. By using AI to simulate structured interview experiences, we created a digital application that enabled autonomous learning, immediate feedback, and measurable improvement in spoken English performance.

Context and Rationale

The Professional Spoken Communication course (LSC 1503) emphasizes real-world speaking skills, including interview performance, clarity, fluency, and confidence. Traditionally, instructors conduct mock interviews and offer feedback based on observed performance.

However, with large class sizes and limited time, providing every learner with meaningful, repeated speaking practice can be difficult. Additionally, many students experience anxiety when speaking English, particularly in formal or high-stakes situations such as job interviews.

Recognizing these limitations, we explored how AI could extend learning opportunities. The goal was to design a tool that allowed students to practice independently, receive structured feedback, and build confidence in a supportive, low-pressure environment.

AI-driven simulations offered an appealing solution: accessible anytime, adaptive to learner input, and capable of providing instant feedback on fluency, vocabulary, and grammar. The customized GPT thus became both a technological and pedagogical innovation—a bridge between classroom teaching and real-world communication practice.

Methodology

Development of the Customized GPT

The project followed a step-by-step approach to designing and integrating the AI tool:

1. **Identifying Learning Objectives:** The main objectives were to help students practice structured interview responses, improve fluency and pronunciation, and gain confidence in professional communication.
2. **Embedding Prompts and Rubrics:** The GPT was programmed to conduct a five-question interview. Two questions were fixed, while three were randomly selected from a pool of twelve, ensuring variation and unpredictability.
3. **Prompt Engineering:** Guided prompts were designed to simulate interviewer follow-ups and provide cues when students struggled to respond.
4. **Real-Time Feedback:** The AI was programmed to give immediate, constructive feedback on fluency, vocabulary, grammar, and overall coherence.
5. **Testing and Refinement:** The tool was piloted with a small group of students, whose feedback informed adjustments to tone, question difficulty, and feedback precision.

Implementation

Over 200 students aged 18–23, at CEFR levels B1–B1+, used the tool during a six-week period as part of their interview preparation. Students accessed the GPT via the ChatGPT interface and engaged in multiple interview simulations. Instructors monitored usage informally and gathered feedback through class discussions, reflections, and post-activity surveys. The focus was on observing changes in engagement, confidence, and communicative performance.

Findings

Student engagement increased notably. Many reported that the AI offered a comfortable space for practice, free from the anxiety of peer judgment. Shy or less confident students in particular appreciated the privacy and flexibility of practicing independently. Frequent users demonstrated marked improvement in fluency and organization of ideas during their final interviews.

Feedback from students emphasized the usefulness of real-time suggestions. The GPT's instant feedback allowed learners to make immediate corrections and understand their linguistic patterns. One student commented that the AI “made me realize my pauses and grammar mistakes right away, so I could fix them next time.” Another shared that “it felt like practicing with a real interviewer, but without pressure.”

Faculty feedback echoed these positive outcomes. Instructors observed more confident performances in final assessments and noted that students were better prepared for spontaneous questions. The tool also encouraged self-directed learning, as students took ownership of their progress outside class time.

However, some limitations emerged. The AI could not assess non-verbal communication skills, such as eye contact or body language, which are crucial in interview contexts. Additionally, initial onboarding required explicit guidance on how to use the tool effectively—students needed orientation to understand the scope and function of AI-driven feedback.

Discussion

This initiative demonstrated how digital applications, when thoughtfully designed, can address persistent challenges in ESL instruction. The customized GPT served as a digital teaching assistant—one that offered unlimited speaking opportunities, consistent feedback, and individualized practice. By leveraging AI, teachers were able to extend learning beyond the classroom and provide authentic, data-informed support.

From a pedagogical standpoint, this project highlights the importance of human-AI collaboration in education. While AI facilitated repetitive practice and immediate feedback, instructors remained central in guiding reflection, contextualizing responses, and reinforcing communication strategies. The success of the project lay not in replacing the teacher, but in empowering both students and instructors through technology.

Moreover, the project underscored the value of adaptability. Students could engage at their own pace and comfort level, and teachers could tailor activities around AI-generated data insights. This flexibility aligns well with contemporary approaches to learner autonomy and differentiated instruction in TESOL.

Pedagogical Implications

The integration of customized GPT tools in ESL programs offers several key takeaways for educators:

- **Scalable Practice:** AI expands opportunities for individualized speaking practice, even in large classes.
- **Instant Feedback:** Real-time evaluation promotes self-awareness and continuous improvement.
- **Reduced Anxiety:** Learners gain confidence through risk-free practice environments.
- **Teacher Support:** Educators can reallocate time toward higher-order tasks, such as analytical feedback and personalized coaching.

However, effective implementation requires careful preparation. Teachers must ensure students understand both the possibilities and limitations of AI. Ethical use, data privacy, and critical awareness of machine feedback should form part of the digital literacy curriculum.

Conclusion and Future Directions

The development and implementation of a customized GPT for ESL interview preparation demonstrates how digital applications can transform traditional language learning into interactive, autonomous, and confidence-building experiences. Students not only improved their fluency and readiness for professional communication but also developed a more positive attitude toward AI as a learning partner.

Looking ahead, the next stage of this project involves exploring **multimodal AI systems** that can assess non-verbal cues such as tone, gesture, and eye contact—elements vital for authentic interview communication. Further research may also focus on **adaptive learning pathways**, where AI tailors practice sessions based on individual learner progress. Such developments could pave the way for a hybrid model of language instruction that combines the empathy of human teaching with the efficiency and personalization of AI.

Keywords: AI in Education, ESL, Interview Preparation, Customized GPT, Digital Learning, TESOL, Professional Communication

Boosting EFL learners' motivation through online project-based learning

Assoc.Prof.Dr. Aydan IRGATOĞLU

Ankara Haci Bayram Veli University, Ankara, Türkiye

APA Citation:

Irgatoğlu, A. (2025). Boosting EFL learners' motivation through online project-based learning

Abstract

Project-based learning (PBL) is widely recognized as a powerful way to engage learners and build motivation in English as a foreign language classrooms. Nevertheless, its impact in fully online settings is still not well understood, even though online classes have become a regular part of education. This study explored how online PBL influences students' motivation, focusing on the L2 Motivational Self System (Dörnyei, 2005), in a preparatory school of a state university. Thirty students took part in a semester-long online PBL course, where they chose topics from Massive Open Online Courses (MOOCs) and presented them to their peers. Using surveys and student reflections, the study found that learners' "ideal L2 self" improved significantly after the course. The flexible and supportive learning environment as self-paced study, freedom to choose topics, and opportunities for peer interaction, played an important role in boosting motivation. These findings suggest that fully online PBL can be an effective way to inspire and sustain student motivation, offering practical ideas for teachers who design online EFL courses.

Keywords: online project-based learning; EFL motivation; L2 self; MOOCs; online teaching practice

Introduction

Project-based learning (PBL) has long been recognized as an effective instructional approach that promotes engagement, autonomy, and motivation among learners of English as a foreign language (EFL). By encouraging students to explore real-world issues, collaborate with peers, and produce meaningful outcomes, PBL supports both linguistic development and personal growth (Stoller, 2006; Thomas, 2017). Traditionally implemented in face-to-face classrooms, this approach has been shown to enhance language proficiency while nurturing learners' sense of agency and intrinsic motivation (Blumenfeld et al., 1991; Park & Hiver, 2017). However, the increasing shift toward online education, accelerated by the COVID-19 pandemic, has transformed how project-based methodologies are applied and perceived in language teaching. Although research on digital and remote language learning has grown considerably (Tao & Gao, 2022; Khong et al., 2023), the motivational impact of fully online PBL environments remains relatively underexplored.

In recent years, online learning has evolved from a temporary alternative to a central component of higher education, especially within EFL programs (Dos Santos, 2022). The flexibility, accessibility, and autonomy afforded by online courses have created new opportunities for learners to take ownership of their education. At the same time, the absence of in-person interaction and immediate feedback from instructors can affect emotional engagement and diminish motivation (Phuong, 2022). Understanding how PBL functions in fully online settings is therefore essential, since it emphasizes self-directed inquiry, collaboration, and authentic communication—elements that may manifest differently in virtual environments.

To investigate motivation in online PBL contexts, this study draws on Dörnyei's (2005) L2 Motivational Self System (L2MSS), which conceptualizes motivation as a dynamic interaction between learners' future self-images and their immediate learning experiences. The ideal L2 self, or the learner's vision of themselves as a proficient user of the target language, is a key predictor of

sustained engagement and effort (Csizér & Kormos, 2014; Dörnyei & Ushioda, 2013). However, how this motivational construct develops in fully online project-based settings remains unclear. Online PBL often allows students to study at their own pace, select topics that are personally relevant, and collaborate virtually with peers, thereby creating conditions that may nurture the ideal L2 self through autonomy, reflection, and meaningful interaction.

Building on these considerations, the present study investigates the effects of fully online PBL on EFL learners' motivation within the framework of the L2 Motivational Self System. The research was conducted with preparatory-school students at a state university who participated in a semester-long online PBL course. Learners selected topics from Massive Open Online Courses (MOOCs), studied them independently, and presented their findings to classmates. By combining quantitative survey data with qualitative reflections, the study aims to provide a nuanced understanding of how online PBL influences learners' motivational development and engagement in digital learning contexts. Specifically, it explores how participation in an online PBL course affects learners' ideal L2 self and motivated learning behavior, and what aspects of the online environment contribute to changes in motivation.

Literature review

2.1. Project-Based Learning in EFL Education

Project-based learning is a learner-centered instructional approach that integrates content knowledge, language practice, and higher-order thinking skills through the completion of authentic, real-world projects (Beckett & Slater, 2005; Stoller, 2006). Unlike traditional teacher-fronted methods, PBL encourages students to take an active role in their learning by engaging in inquiry, collaboration, and problem-solving (Thomas, 2017). In EFL contexts, it has been shown to enhance communicative competence, autonomy, and motivation (Fragoulis, 2009; Le & Nguyen, 2022). Through meaningful tasks such as research projects, oral presentations, and multimedia productions, learners use English as a medium to construct and communicate knowledge, reinforcing both linguistic and cognitive growth (Beckett, 2002).

A substantial body of research supports the benefits of PBL for language learning. Studies have found that project-based tasks foster engagement, promote cooperative learning, and encourage a sense of ownership over the learning process (Chu et al., 2016; Guo et al., 2020). PBL is closely aligned with constructivist theories of learning, as it enables students to build understanding through exploration, reflection, and social interaction (Dewey, 1938; Vygotsky, 1978). Motivation consistently emerges as one of its strongest outcomes, since students often perceive their projects as personally relevant and experience satisfaction and pride as they complete them (Simpson, 2011; Nguyen & Nguyen, 2023).

2.2. Online Learning and PBL in Digital Contexts

The rapid expansion of digital education has led to the adaptation of PBL to online and hybrid contexts, allowing learners to collaborate virtually and engage in authentic communication through technology (Mishra & Koehler, 2006; Krajcik & Shin, 2014). Online PBL typically utilizes learning management systems, video conferencing, and collaborative digital workspaces that facilitate research, idea exchange, and virtual project presentations (Chen & Yang, 2019). While these tools provide flexibility and accessibility, they also reshape classroom dynamics and require new pedagogical approaches.

Online environments support learner autonomy and self-regulation, two critical aspects of motivation (Broadbent & Fuller-Tyszkiewicz, 2018; Sun & Chen, 2016). However, the lack of face-to-face interaction may reduce learners' sense of connectedness and affective engagement (Phuong, 2022). In fully online PBL, maintaining a balance between independence and social interaction becomes

particularly important. Learners benefit from the freedom to select project topics that reflect their interests, yet they also require structured opportunities for peer collaboration and feedback to sustain motivation. Empirical studies have shown that when online PBL is well-designed, scaffolded, and interactive, it can enhance engagement, self-efficacy, and learning outcomes (Hernández-Ramos & De La Paz, 2009; Khong et al., 2023). Nevertheless, evidence on the motivational processes underlying these outcomes, particularly in EFL settings, remains limited.

2.3. L2 Motivation and the Motivational Self System

Motivation is widely recognized as a key determinant of success in second and foreign language learning (Gardner, 1985; Dörnyei, 2001). Among the most influential frameworks in this area is Dörnyei's (2005) L2 Motivational Self System, which reconceptualizes motivation in relation to self-identity. The framework comprises three interconnected components: the ideal L2 self, the ought-to L2 self, and the L2 learning experience. The ideal L2 self reflects the learner's vision of themselves as a competent user of the target language, and it serves as a powerful internal motivator that drives sustained effort (Dörnyei & Ushioda, 2013). The ought-to L2 self represents perceived obligations and external expectations, while the L2 learning experience refers to immediate attitudes toward the learning environment and tasks.

Research across diverse contexts has consistently shown that the ideal L2 self is a strong predictor of motivated learning behavior (Papi & Khajavy, 2021; Taguchi et al., 2009). Learners who can vividly imagine themselves as successful language users tend to demonstrate greater persistence and engagement. Furthermore, positive learning experiences such as supportive teaching, autonomy, and authentic communication strengthen the connection between learners' future self-guides and their current learning actions (You & Dörnyei, 2016). Creating learning environments that activate and maintain the ideal L2 self is therefore essential to sustaining long-term motivation.

The theoretical principles of PBL and the L2 Motivational Self System converge in meaningful ways. PBL's focus on autonomy, collaboration, and authentic communication directly fosters the development of the ideal L2 self by helping learners visualize themselves as competent users of English in real-world contexts (Henry, 2015). Allowing students to choose topics that reflect their interests fosters a sense of personal relevance and ownership, reinforcing their future self-concept. The social dimension of PBL, particularly through peer feedback and public presentation, enhances learners' confidence and contributes to a richer L2 learning experience (Busse, 2013; Kormos & Csizér, 2014).

In online contexts, these mechanisms may operate differently. The flexibility of online PBL supports autonomy and self-paced learning, while digital platforms extend interaction beyond classroom boundaries and offer diverse modes of communication. Such environments can nurture learners' ideal L2 self by providing authentic, multimodal opportunities to practice English and express their identities. However, the motivational potential of online PBL depends on achieving an appropriate balance between independence and social engagement to ensure sustained interest and effort throughout the learning process.

Although the motivational impact of PBL has been widely documented in traditional classroom settings, little is known about how it operates in fully online EFL environments. The present study seeks to fill this gap by examining the influence of a semester-long online PBL course on learners' motivation within the framework of the L2 Motivational Self System. Conducted with preparatory school students who selected topics from MOOCs, studied them independently, and shared their findings with peers, the study aims to explore how autonomy, flexibility, and collaboration in online PBL settings contribute to changes in learners' ideal L2 self and overall motivation. The findings are expected to extend theoretical understanding of motivation in digital contexts while offering practical insights for designing engaging and learner-centered online EFL instruction.

Method

This study adopted a mixed-method research design to examine how participation in a fully online project-based learning (PBL) course influenced EFL learners' motivation within the framework of the L2 Motivational Self System. Quantitative and qualitative data were collected to capture both measurable changes and learners' subjective experiences.

3.1. Participants

The participants were thirty intermediate-level students enrolled in the English preparatory program of a state university in Türkiye. Their ages ranged from eighteen to twenty-one, and all had studied English for at least four years prior to the course. None of the students had previous experience with project-based learning in an online setting. Participation was voluntary, and informed consent was obtained from all students before data collection began. To ensure relative homogeneity, the participants' English proficiency levels were verified through institutional placement test results administered at the beginning of the semester. All students were placed in the same intact class, which followed a ten-week online PBL-based English communication course.

3.2. Context and Course Design

The study was integrated into a ten-week online English communication skills course offered through the university's learning management system (LMS). The course followed a project-based model emphasizing student collaboration, inquiry, and authentic communication. Each project cycle consisted of three phases: topic selection, research and preparation, and online presentation.

Students were encouraged to select topics of personal interest from Massive Open Online Courses (MOOCs) available on platforms such as Coursera and edX. These open-access resources provided authentic, content-rich materials in English that aligned with students' academic and personal interests. Learners explored these materials independently, synthesized relevant information, and prepared short digital presentations or projects to share during synchronous online sessions. The instructor acted as a facilitator throughout the process, offering guidance on research organization, language use, and presentation design. Students collaborated using digital tools such as Google Docs, Zoom breakout rooms, and LMS discussion forums. This structure was intended to promote autonomy, self-paced learning, and peer interaction, which are central elements of both PBL and the development of the ideal L2 self.

3.3. Instruments

Data were collected through two instruments: a Motivational Self System Questionnaire and Student Reflection Reports.

The Motivational Self System Questionnaire was adapted from Dörnyei (2005) and You and Dörnyei (2016) and consisted of twenty-four items measuring three subscales: ideal L2 self, ought-to L2 self, and L2 learning experience. Each item was rated on a six-point Likert scale ranging from one (strongly disagree) to six (strongly agree). The scale demonstrated strong internal consistency in this study (Cronbach's $\alpha = .89$).

The Student Reflection Reports were written narratives in which participants described how the online PBL experience affected their motivation, confidence, and engagement with English learning. These reflections provided rich qualitative data that complemented the quantitative findings and offered deeper insight into learners' motivational development.

The study followed a one-group pretest–posttest mixed-method design. At the beginning of the semester, participants completed the Motivational Self System Questionnaire to establish baseline measures of their motivation. Over the course of ten weeks, they participated in three major project cycles involving topic selection, research using MOOC materials, and presentation of their findings in English during online class sessions. After each presentation, students received feedback from both their peers and the instructor. Reflection tasks were also integrated into the lessons to encourage self-assessment and goal setting. At the end of the course, participants completed the same questionnaire as a posttest and submitted their written reflections.

3.4. Data Analysis

Quantitative data were analyzed using SPSS 26. Descriptive statistics were calculated to summarize means and standard deviations for each subscale of the L2 Motivational Self System. Paired-samples t-tests were used to determine whether differences between pretest and posttest scores were statistically significant. Effect sizes (Cohen's d) were also computed to assess the magnitude of observed changes.

Qualitative data were analyzed through thematic analysis following Braun and Clarke (2006). Student reflections were read multiple times to identify recurring patterns and themes related to autonomy, engagement, peer collaboration, and motivational change. Codes were grouped into broader categories and cross-referenced with quantitative results to achieve data triangulation and ensure a comprehensive interpretation of the findings.

1. Results and Discussion

4.1. Quantitative Findings

The quantitative results indicated a significant improvement in students' motivation after completing the online PBL course. The mean pretest and posttest scores for each subscale of the L2 Motivational Self System are presented in Table 1.

Table 1. The mean pretest and posttest scores for each subscale

Subscale	Pretest M (SD)	Posttest M (SD)	t(29)	p	Cohen's d
Ideal L2 Self	4.21 (0.56)	5.01 (0.48)	6.43	< .001	1.05
Ought-to L2 Self	4.02 (0.67)	4.11 (0.60)	0.72	.48	0.09
L2 Learning Experience	4.35 (0.63)	4.92 (0.52)	4.18	< .001	0.76

There was a statistically significant increase in learners' ideal L2 self and L2 learning experience scores, with large and moderate effect sizes respectively. This finding suggests that the online PBL environment enhanced both learners' long-term motivational vision and their immediate engagement with English learning. In contrast, the change in the ought-to L2 self was not statistically significant, indicating that students' perceptions of external expectations or obligations remained stable. The overall results imply that the motivational gains observed were primarily internally driven and emerged from learners' enjoyment and active participation in the online PBL process.

4.2. Qualitative Insights

The qualitative analysis revealed three interrelated themes that illuminate the motivational processes underlying these quantitative findings: autonomy and self-direction, peer interaction and community building, and personal relevance and confidence.

Students frequently emphasized that the freedom to choose topics played a crucial role in sustaining their interest and effort. Selecting MOOCs that matched their personal goals or curiosities gave the learning process a sense of authenticity and meaning. One participant expressed this sentiment by stating, “For the first time, I studied English because I wanted to, not because I had to. I felt that what I was learning had a purpose.” This sense of ownership reflects the activation of the ideal L2 self, as learners began to view themselves as self-determined English users capable of managing their learning.

Peer interaction and support also emerged as a vital element of motivation. Despite the virtual format, students reported feeling connected to their classmates through group discussions, collaborative work, and presentation feedback sessions. One learner commented, “We were far away from each other, but teamwork made me feel part of something real.” These experiences resonate with Dörnyei’s (2005) notion of the L2 learning experience as a situational component of motivation shaped by classroom dynamics and social engagement.

Confidence and identity growth formed the third major theme. Many students explained that presenting their projects in English helped them overcome self-doubt and strengthened their belief in their communicative abilities. One reflection captured this transformation: “I started to see myself as someone who can communicate in English confidently. It changed how I see my future self.” This emerging sense of self-assurance demonstrates how authentic communicative experiences within PBL can reinforce learners’ vision of themselves as capable language users.

Taken together, the quantitative and qualitative results suggest that online PBL effectively enhances motivation by promoting autonomy, meaningful interaction, and emotional engagement—key components of the ideal L2 self and the L2 learning experience. The increase in quantitative scores reflects learners’ cognitive and emotional investment, while the qualitative narratives reveal the mechanisms driving this change, including personal relevance, peer collaboration, and reflective learning.

These findings are consistent with earlier research demonstrating that PBL promotes motivation by linking language learning to personally meaningful goals (Guo et al., 2020; Le & Nguyen, 2022). However, the current study extends prior knowledge by confirming that these motivational benefits can also emerge in fully online environments. The flexibility of the online setting allowed learners to study independently and pursue individualized interests, while structured interaction maintained a sense of community and shared purpose. This balance of autonomy and social engagement appears to be crucial for sustaining motivation in remote learning contexts, echoing the observations of Broadbent and Fuller-Tyszkiewicz (2018) and Phuong (2022).

The findings highlight several important implications for online EFL instruction. Allowing learners to choose project topics that align with their interests enhances intrinsic motivation and personal investment in the learning process. Structuring online activities that encourage collaboration and peer feedback helps preserve the sense of social connectedness that is often missing in virtual classrooms. Furthermore, integrating authentic digital materials such as MOOCs and open educational resources exposes learners to meaningful language use and reinforces their identity as English users in global contexts.

Teachers play a pivotal role in balancing learner autonomy with instructional guidance. By acting as facilitators and motivational mentors, they can help students articulate, visualize, and strengthen their ideal L2 selves. Encouraging reflective practices—such as journaling, goal setting, and self-assessment—further reinforces the link between learners’ current actions and their envisioned future selves, thereby supporting sustained motivation beyond the course.

Overall, the results of this study demonstrate that fully online project-based learning can significantly enhance learners’ motivation by fostering autonomy, social interaction, and authentic engagement.

Through meaningful project work, students develop a clearer vision of themselves as capable L2 users and experience greater satisfaction and confidence in their learning. Although the relatively small sample size limits the generalizability of the findings, the study provides promising evidence that online PBL can be an effective and motivating pedagogical approach for EFL learners in digital education contexts.

2. Conclusions

The present study examined the impact of fully online project-based learning (PBL) on EFL learners' motivation within the framework of the L2 Motivational Self System (L2MSS). By combining quantitative and qualitative data from a ten-week online communication course, the study sought to understand how features of online PBL—such as autonomy, flexibility, and peer collaboration—contribute to the development of learners' ideal L2 self and overall motivation. The findings revealed significant increases in students' ideal L2 self and L2 learning experience, supported by qualitative evidence of heightened self-direction, peer engagement, and confidence. These results suggest that when effectively structured, online PBL can foster meaningful and sustainable motivational growth among EFL learners.

The quantitative results demonstrated that students developed stronger motivational self-images and more positive attitudes toward English learning after participating in the online PBL course. The qualitative reflections provided complementary insights, showing that the opportunity to choose personally relevant topics, collaborate with peers, and engage in authentic communication encouraged learners to perceive English not merely as an academic subject but as a valuable tool for self-expression and connection. The combination of independent inquiry and social collaboration enabled students to internalize motivation and construct a more vivid and attainable vision of their ideal L2 self.

These findings have several pedagogical and theoretical implications. From a pedagogical perspective, online PBL offers a viable model for enhancing learner engagement and motivation in digital EFL contexts. By allowing students to make choices, work collaboratively, and connect learning tasks to authentic goals, teachers can create online environments that nurture autonomy and sustained interest. The integration of authentic materials such as MOOCs or open educational resources can further enrich these experiences, providing exposure to diverse language input and global perspectives. Instructors should also incorporate structured reflection activities that prompt learners to visualize their ideal L2 self and link their ongoing efforts to this future vision. Such reflective components transform PBL from a task-based activity into a motivationally rich experience grounded in self-regulation and identity development.

Theoretically, the study extends the application of the L2 Motivational Self System by demonstrating its relevance in fully online, project-based contexts. The findings underscore that motivational self-guides are not static constructs but can be actively shaped through pedagogical design. When learners are given opportunities to experience autonomy, meaningful interaction, and success in authentic communication, their ideal L2 self becomes more salient and attainable. This interplay between online learning design and motivational self-construction highlights the importance of aligning digital pedagogies with principles of learner identity and vision building.

Despite its promising results, the study has certain limitations. The small sample size and single-institution context limit the generalizability of the findings. Future research could replicate the study with larger and more diverse populations to validate the results and explore potential variations across proficiency levels, cultural backgrounds, or instructional settings. Longitudinal research would also be valuable to examine whether the motivational gains observed in online PBL are sustained over time. Additionally, future studies might investigate how specific design features—such as the degree of scaffolding, peer feedback mechanisms, or types of digital tools—affect different components of the L2 Motivational Self System.

In conclusion, this study provides empirical support for the integration of project-based learning into fully online EFL programs. By combining flexibility with structured collaboration and reflection, online PBL can cultivate not only linguistic competence but also deeper motivational engagement. When learners are encouraged to take ownership of their learning, connect it to personally meaningful goals, and envision themselves as confident English users, their motivation evolves from external compliance to internalized commitment. As online education continues to expand globally, designing digital courses that foster such motivational transformation is both a pedagogical necessity and an opportunity to empower language learners as autonomous, self-determined individuals.

References

Abbas, A., Ashiq, M., & Asghar, A. (2019). The impact of digital literacy on students' academic performance at the university level. *Bulletin of Education and Research*, 41(2), 65–78.

Amiri, F. (2009). The relationship between computer literacy and EFL students' academic achievement. *Asian EFL Journal*, 11(3), 120–135.

Belda-Medina, J. R. (2022). Digital literacy and autonomous learning in EFL contexts: A systematic review. *Computer Assisted Language Learning*, 35(5–6), 1284–1304. <https://doi.org/10.1080/09588221.2020.1839501>

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3–4), 369–398. <https://doi.org/10.1080/00461520.1991.9653139>

Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2019). Digital transformation in higher education: A framework for digital literacy development. *Educational Technology Research and Development*, 67(5), 1235–1267. <https://doi.org/10.1007/s11423-019-09641-2>

Broadbent, J., & Fuller-Tyszkiewicz, M. (2018). Profiles in self-regulated learning and their correlates for online and blended learning students. *Educational Technology Research and Development*, 66(6), 1435–1455. <https://doi.org/10.1007/s11423-018-9595-9>

Brown, C. (2009). Digital literacy and language learning: Expanding definitions and practices. *TESOL Quarterly*, 43(4), 731–734. <https://doi.org/10.1002/j.1545-7249.2009.tb00191.x>

Carolus, A., Binder, J. F., Muench, R., Schmidt, C., & Schneider, F. (2023). Conceptualizing digital literacy in education: Beyond technical competence. *Computers & Education*, 194, 104654. <https://doi.org/10.1016/j.compedu.2022.104654>

Csizér, K., & Kormos, J. (2014). The L2 motivational self system: A review of empirical studies. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 71–89). Palgrave Macmillan.

Dewaele, J. M., & Li, C. (2021). Emotions in second language acquisition: A critical review and research agenda. *Foreign Language World*, 203(1), 34–49.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.

Dörnyei, Z., & Ushioda, E. (2013). *Teaching and researching motivation* (2nd ed.). Routledge.

Dos Santos, L. M. (2022). Online and hybrid EFL learning environments: A review of challenges and practices. *Education and Information Technologies*, 27(1), 889–905. <https://doi.org/10.1007/s10639-021-10636-2>

Fredrickson, B. L. (2003). The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society B*, 358(1431), 1367–1377. <https://doi.org/10.1098/rstb.2003.1399>

Guo, Y., Xu, J., & Li, C. (2020). Project-based learning and learner motivation in EFL classrooms: A longitudinal study. *Innovation in Language Learning and Teaching*, 14(5), 427–440. <https://doi.org/10.1080/17501229.2019.1615490>

Hafner, C. A. (2014). Embedding digital literacies in English language teaching: Students' digital video projects. *TESOL Quarterly*, 48(4), 655–685. <https://doi.org/10.1002/tesq.138>

Hockly, N. (2015). Developments in digital literacies. *ELT Journal*, 69(3), 308–313. <https://doi.org/10.1093/elt/ccv020>

Khong, T. D. H., Nguyen, L. H., & Le, P. T. (2023). Online project-based learning and EFL learners' engagement: Voices from Vietnamese university students. *Language Learning & Technology*, 27(1), 1–21. <https://doi.org/10.1016/j.llt.2022.03.004>

Le, T. T. H., & Nguyen, T. M. H. (2022). Project-based learning and motivation in online EFL contexts: A mixed-methods study. *Asian EFL Journal*, 26(2), 98–122.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283–305. <https://doi.org/10.1111/j.1467-1770.1994.tb01103.x>

Mudra, H. (2020). Digital literacy among EFL students and its relationship with language learning outcomes. *Teaching English with Technology*, 20(3), 26–39.

Palilonis, J., & Watt, J. (2019). Rethinking digital literacy in the 21st century. *Journal of Media Literacy Education*, 11(3), 57–69. <https://doi.org/10.23860/JMLE-2019-11-03-06>

Park, G. P., & Hiver, P. (2017). Profiling classroom motivation: The role of the L2 motivational self system in project-based learning. *Language Teaching Research*, 21(6), 719–739. <https://doi.org/10.1177/1362168816684368>

Phuong, H. N. (2022). Motivation and engagement in online English learning during COVID-19: A student perspective. *Computer Assisted Language Learning Electronic Journal*, 23(2), 114–134.

Stoller, F. (2006). Establishing a theoretical foundation for project-based learning in second and foreign language contexts. In G. H. Beckett & P. C. Miller (Eds.), *Project-based second and foreign language education: Past, present, and future* (pp. 19–40). Information Age Publishing.

Tao, J., & Gao, X. (2022). Language teacher agency and online teaching during the COVID-19 pandemic: The case of project-based English instruction. *System*, 105, 102718. <https://doi.org/10.1016/j.system.2022.102718>

Thomas, J. W. (2017). *A review of research on project-based learning*. Buck Institute for Education.

Wang, Y. (2023). Digital literacy, language learning, and technology integration in the post-pandemic classroom. *Computer Assisted Language Learning*, 36(7), 1023–1040. <https://doi.org/10.1080/09588221.2022.2064354>

AUTHOR BIODATA

Dr. Aydan Irgatoğlu works as an associate professor at the School of Foreign Languages of Ankara Hacı Bayram Veli University. She is also working there as the president. She is a graduate of the English Language Teaching Department of Hacettepe University. She completed her MA and Ph.D. at the same department of Gazi University. She worked at Başkent University for 10 years and she also worked as the head of the Department of Translation and Interpreting at Başkent University. Her interest areas are L2 teacher education and language learning strategies.

How digital literacies shape EFL learning: Reducing stress, easing anxiety, and boosting enjoyment

Assoc. Prof. Dr. Aydan IRGATOĞLU

Ankara Hacı Bayram Veli University, Ankara, Türkiye

APA Citation:

Irgatoglu, A. (2025). How digital literacies shape EFL learning: Reducing stress, easing anxiety, and boosting enjoyment.

Abstract

Learning English as a foreign language can be both exciting and stressful. While many learners enjoy the process, others struggle with anxiety and technostress, which can affect their progress. This study examined how developing digital literacy skills influences learners' emotions and achievements. 122 intermediate EFL students studying at the preparatory school of a state university were divided into two groups: one was taught with traditional methods, and the other used technology-based lessons with a focus on digital literacy. Before and after the course, students were tested on technostress, foreign language anxiety, academic enjoyment, and language achievement. Results showed that students who learned with digital tools experienced less anxiety and technostress, more enjoyment, and higher achievement than those in the traditional group. These findings suggest that digital literacies not only improve language learning but also create a more positive and supportive learning experience. For teachers, integrating digital tools into the classroom can help students feel more confident, motivated, and successful in their language learning journey

Keywords: technostress; academic enjoyment; digital literacies; language anxiety

Introduction

In today's digitally connected world, technology has become an indispensable component of education, transforming the ways in which languages are taught and learned. The integration of digital tools into English as a foreign language (EFL) classrooms has created both opportunities and challenges for learners and teachers. On one hand, technology makes language learning more interactive, accessible, and motivating; on the other, it can provoke anxiety and technostress among students who struggle to adapt to digital environments (Tarañdar et al., 2019; Wang, 2023). Consequently, learning English as a foreign language can be both an exciting and a stressful experience, depending on learners' levels of technological competence and emotional resilience.

In this context, digital literacy, defined as the ability to access, evaluate, and effectively use digital resources, has emerged as a key competence for success in twenty-first-century education (Heidari & Tabatabaei-Yazdi, 2021; Pérez-Escoda et al., 2019). Students who possess higher levels of digital literacy are better equipped to navigate online platforms, manage digital tools, and engage confidently with multimodal materials. Conversely, limited digital literacy skills may heighten technostress and hinder meaningful engagement in language learning activities (Joo et al., 2016; Verkijika, 2019). As EFL classrooms increasingly depend on technology, helping learners develop digital literacy has become essential for fostering both effective learning experiences and positive emotional responses.

At the same time, the affective dimension of learning, which encompasses emotions such as anxiety, enjoyment, and motivation, plays a crucial role in determining learners' success in acquiring a foreign language (Dewaele & MacIntyre, 2014; Pekrun & Perry, 2014). Negative emotions such as anxiety and technostress can undermine self-confidence and academic performance, whereas positive emotions like enjoyment can strengthen engagement and achievement (Dewaele & Li, 2021; Pinel & Albert, 2018). Despite increasing recognition of the affective dimension, the interplay between digital

literacy and learners' emotions remains underexplored in EFL contexts, particularly in higher education, where students are expected to use technology extensively.

The present study investigates how developing digital literacy skills influences EFL learners' emotional and academic outcomes. It examines the effects of technology-based instruction emphasizing digital literacy on technostress, foreign language anxiety, academic enjoyment, and language achievement. By comparing traditional and digitally enhanced learning environments among intermediate university-level EFL learners, this study aims to provide empirical evidence on how the integration of digital literacy into language instruction can foster higher academic achievement alongside greater emotional well-being and confidence in learning. Specifically, the study addresses four research questions: whether developing digital literacy reduces technostress, lowers foreign language anxiety, increases academic enjoyment, and enhances language achievement among EFL learners.

Literature review

The rapid digitalization of education has reshaped how knowledge is constructed, shared, and communicated in language learning contexts. Digital literacy extends beyond technical competence to include the ability to create, interpret, and communicate meaning using digital tools within socially and culturally situated contexts (Palilonis & Watt, 2019; Hafner, 2014). Initially conceptualized by Zirkowski (1974) as the capacity to recognize, locate, and evaluate information, digital literacy has evolved to include multiple dimensions such as media, information, and computer literacy (Carolus et al., 2023; Wang, 2023). In EFL classrooms, digital literacy enables learners to interact with authentic materials, collaborate online, and develop communicative competence through multimodal platforms (Hafner, Chik, & Jones, 2015).

A growing body of research emphasizes the pedagogical value of digital literacy for both learners and teachers. Digitally literate students demonstrate higher motivation, greater autonomy, and more positive attitudes toward language learning (Belda-Medina, 2022; Hockly, 2015). Teachers, in turn, can leverage digital tools to create interactive and personalized learning environments that encourage learner participation (Gudmundsdottir & Hatlevik, 2018). Nevertheless, many educators and students continue to face challenges in applying digital literacy effectively, often due to insufficient training or unequal access to technological resources, a situation commonly referred to as the digital divide (Bond et al., 2019; Palacios-Hidalgo & Huertas-Abril, 2022).

While technology improves instructional efficiency, it can also produce psychological strain. Technostress, defined as the stress or anxiety resulting from an inability to cope with technological demands, has been identified as a growing concern in educational contexts (Taraifdar et al., 2019). Learners with limited digital competence are more likely to experience technostress, which can reduce their willingness to engage with technology in learning environments (Joo et al., 2016; Verkijika, 2019). In EFL settings, technostress may manifest as frustration, cognitive overload, or avoidance of digital tools, all of which can hinder concentration and engagement (Maier et al., 2019; Steelman & Soror, 2017).

Technostress frequently interacts with foreign language anxiety (FLA), defined as the apprehension learners experience when using the target language (Horwitz et al., 1986; MacIntyre & Gardner, 1994). Prior studies indicate that both technology-related pressure and language anxiety can negatively affect learners' self-efficacy and participation (Hafner & Ho, 2020; Han et al., 2022). However, research also shows that structured digital literacy instruction can help reduce anxiety by improving learners' confidence in managing both linguistic and technological demands (Dashtestani & Hojatpanah, 2020; Rezai et al., 2024).

Positive emotions, particularly enjoyment, have also gained attention in recent language learning research. According to Fredrickson's (2003) Broaden-and-Build Theory, positive emotions broaden individuals' thought-action repertoires and build enduring personal resources. In second language acquisition, foreign language enjoyment (FLE) represents a vital emotional factor that enhances persistence, collaboration, and achievement (Dewaele & Li, 2021; Pinel & Albert, 2018). Empirical evidence suggests that enjoyment and anxiety often coexist yet function as distinct predictors of learning outcomes (Li, 2021; Zeng, 2021). Learners who find digital learning activities enjoyable are more engaged and perform better than those who perceive them as stressful (Gregersen, MacIntyre, & Meza, 2014). Hence, enhancing digital literacy may serve as a pathway to emotional balance by simultaneously reducing technostress and promoting enjoyment.

Although previous studies have demonstrated that digital literacy contributes to motivation, communication, and achievement (Amiri, 2009; Brown, 2009; Fairlie et al., 2010; Abbas et al., 2019; Mudra, 2020), few have explored how it simultaneously affects technostress, foreign language anxiety, academic enjoyment, and achievement, particularly in emerging EFL contexts. Addressing this gap, the present study investigates whether improving digital literacy can reduce learners' negative emotions, increase enjoyment, and enhance academic performance in English language learning.

Method

The participants of the study were 122 intermediate-level students enrolled in the preparatory English program of a state university. Their ages ranged from 18 to 22, and all had studied English for at least four years before the study. The Oxford Quick Placement Test (OQPT) was administered to ensure language proficiency homogeneity, and only students within the intermediate band were included. The participants were randomly assigned to two groups of equal size, forming an experimental group and a control group, each consisting of 61 students. All participants provided informed consent and were assured of confidentiality and voluntary participation.

Four instruments were employed to collect data before and after the instructional intervention. The Technostress Scale developed by Wang and Chee Tan (2022) assessed students' emotional and cognitive stress associated with technology use through 13 Likert-type items. The Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986; Jiang & Dewaele, 2019) measured learners' self-perceived anxiety in English classes through eight items. The Academic Enjoyment Scale, adapted from Jiang and Dewaele (2019), evaluated students' positive engagement in ten items related to enjoyment in private and social dimensions. Finally, a Language Achievement Test consisting of 30 items covering grammar, vocabulary, and reading comprehension was developed based on course materials. All instruments demonstrated acceptable reliability, with Cronbach's alpha coefficients ranging from .84 to .88.

A quasi-experimental pretest–posttest control group design was employed over a 10-week semester. Both groups followed the same English syllabus and used the same instructional materials, but the mode of instruction differed. The experimental group received technology-enhanced instruction designed to improve digital literacy through multimedia resources, online collaboration, and interactive tasks that required students to locate, evaluate, and create digital content in English. The control group was taught using traditional face-to-face methods that relied primarily on textbooks and teacher explanations without explicit digital literacy integration. Before and after the intervention, all participants completed the four measures.

Data were analyzed using SPSS 26. Descriptive statistics were calculated to compare pretest and posttest means. To evaluate the impact of digital literacy instruction, one-way ANCOVA was performed for each dependent variable, controlling for pretest scores. The significance level was set at $p < .05$.

3. Results and Discussion

Pretest analyses revealed no significant differences between groups, confirming baseline equivalence. Posttest results, however, indicated substantial improvements in the experimental group, which reported lower levels of technostress and anxiety and higher levels of enjoyment and language achievement compared to the control group. ANCOVA results confirmed significant effects of instructional method on all four variables: technostress ($F[1,119] = 18.72, p < .001$), foreign language anxiety ($F[1,119] = 14.95, p < .001$), academic enjoyment ($F[1,119] = 23.81, p < .001$), and language achievement ($F[1,119] = 19.44, p < .001$).

The results clearly demonstrate that digital literacy instruction can positively influence both affective and academic aspects of EFL learning. Students exposed to technology-based instruction experienced lower technostress and anxiety while showing higher enjoyment and achievement, supporting earlier research linking digital competence to confidence and engagement (Amiri, 2009; Belda-Medina, 2022). The reduction in technostress aligns with findings suggesting that familiarity with digital tools transforms technology from a stressor into an empowering resource (Joo et al., 2016; Tarafdar et al., 2019). The decrease in foreign language anxiety similarly confirms that digital literacy can enhance learners' sense of control and emotional comfort (Hafner & Ho, 2020).

The increase in enjoyment among the digitally trained learners corresponds with Fredrickson's (2003) Broaden-and-Build Theory, which emphasizes that positive emotions expand learners' cognitive and social resources. Enjoyment encouraged greater engagement, collaboration, and perseverance, ultimately contributing to improved language achievement, consistent with the findings of Dewaele and Li (2021) and Fairlie et al. (2010).

4. Conclusions

This study explored the influence of digital literacy instruction on technostress, foreign language anxiety, academic enjoyment, and achievement among EFL learners. The findings demonstrated that developing digital literacy not only enhances language performance but also promotes emotional well-being by reducing anxiety and stress while fostering enjoyment. These results suggest that digital literacy functions as both a cognitive and affective asset in the foreign language classroom, enabling learners to manage technological demands effectively while deriving satisfaction from the learning process.

From a pedagogical perspective, digital literacy should be systematically integrated into EFL curricula. Teachers need to guide learners in using digital tools strategically, provide emotional support during technological adaptation, and design interactive, creative, and collaborative digital tasks that promote both engagement and enjoyment. Professional development programs should emphasize digital pedagogy and strategies to foster emotional resilience, while institutions must ensure equitable access to technological resources.

Future studies could extend these findings by exploring long-term effects of digital literacy on motivation, resilience, and learner identity across varied educational contexts. Combining quantitative and qualitative approaches, such as interviews and observations, would yield a richer understanding of how digital literacy shapes both the cognitive and affective dimensions of language learning. Ultimately, digital literacy emerges not only as a practical necessity but also as a psychological enabler of confident, autonomous, and enjoyable language learning experiences.

References

Abbas, Q., Hussain, S., & Rasool, S. (2019). Digital literacy effect on the academic performance of students at higher education level in Pakistan. *Global Social Sciences Review*, 5(1), 154–165.

Amiri, S. (2009). The effects of information and communication technology on at-risk children of low economic status: Make It–Take It after-school case study. *International Journal of Education and Development Using Information and Communication Technology*, 5(3), 141–147.

Belda-Medina, J. (2022). Promoting inclusiveness, creativity and critical thinking through digital storytelling among EFL teacher candidates. *International Journal of Inclusive Education*, 26(1), 109–123. <https://doi.org/10.1080/13603116.2021.2011440>

Bond, M., Zawacki-Richter, O., & Nichols, M. (2019). Revisiting five decades of educational technology research: A content and authorship analysis of the *British Journal of Educational Technology*. *British Journal of Educational Technology*, 50(1), 12–63. <https://doi.org/10.1111/bjet.12730>

Brown, B. C. (2009). *An examination of the relationship between digital literacy and student achievement in Texas elementary schools* (Doctoral dissertation, The University of Oklahoma). ProQuest Dissertations Publishing.

Carolus, A., Augustin, Y., Markus, A., & Wienrich, C. (2023). Digital interaction literacy model: Conceptualizing competencies for literate interactions with voice-based AI systems. *Computers and Education: Artificial Intelligence*, 4, 100114. <https://doi.org/10.1016/j.caeari.2022.100114>

Dashtestani, R., & Hojatpanah, S. (2020). Digital literacy of EFL students in a junior high school in Iran: Voices of teachers, students and ministry directors. *Computer Assisted Language Learning*, 35(4), 635–665. <https://doi.org/10.1080/09588221.2020.1744664>

Dewaele, J. M., & Li, C. (2021). Teacher enthusiasm and students' social-behavioral learning engagement: The mediating role of student enjoyment and boredom in Chinese EFL classes. *Language Teaching Research*, 25(6), 922–945. <https://doi.org/10.1177/13621688211014538>

Dewaele, J. M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237–274. <https://doi.org/10.14746/ssllt.2014.4.2.5>

Fairlie, R. W., Beltran, D. O., & Das, K. (2010). Home computers and educational outcomes: Evidence from the NLSY97 and CPS. *Economic Inquiry*, 48(3), 771–792. <https://doi.org/10.1111/j.1465-7295.2009.00163.x>

Fredrickson, B. L. (2003). The value of positive emotions. *American Scientist*, 91(4), 330–335. <https://doi.org/10.1511/2003.4.330>

Gudmundsdottir, G. B., & Hatlevik, O. E. (2018). Newly qualified teachers' professional digital competence: Implications for teacher education. *European Journal of Teacher Education*, 41(2), 214–231. <https://doi.org/10.1080/02619768.2017.1416085>

Hafner, C. A. (2014). Embedding digital literacies in English language teaching: Students' digital video projects as multimodal ensembles. *TESOL Quarterly*, 48(4), 655–685. <https://doi.org/10.1002/tesq.138>

Hafner, C. A., Chik, A., & Jones, R. H. (2015). Digital literacies and language learning. *Language Learning & Technology*, 19(3), 1–7.

Hafner, C. A., & Ho, W. Y. J. (2020). Assessing digital multimodal composing in second language writing: Toward a process-based model. *Journal of Second Language Writing*, 47, 100710. <https://doi.org/10.1016/j.jslw.2020.100710>

Han, S., Li, Y., & Haider, S. A. (2022). Impact of foreign language classroom anxiety on higher education students' academic success: Mediating role of emotional intelligence and moderating influence of classroom environment. *Frontiers in Psychology*, 13, 945062. <https://doi.org/10.3389/fpsyg.2022.945062>

Heidari, N., & Tabatabaee-Yazdi, M. (2021). Digital literacy skills among Iranian EFL teachers and students. *Journal of Research in Techno-based Language Education*, 1(1), 22–34.

Hockly, N. (2015). Developments in online language learning. *ELT Journal*, 69(3), 308–313. <https://doi.org/10.1093/elt/ccv020>

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>

Jiang, Y., & Dewaele, J. M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System*, 82, 13–25. <https://doi.org/10.1016/j.system.2019.02.017>

Joo, Y. J., Lim, K. Y., & Kim, N. H. (2016). The effects of secondary teachers' technostress on the intention to use technology in South Korea. *Computers & Education*, 95(1), 114–122. <https://doi.org/10.1016/j.compedu.2015.12.004>

Li, C. (2021). A control–value theory approach to boredom in English classes among university students in China. *The Modern Language Journal*, 105(2), 317–334. <https://doi.org/10.1111/modl.12693>

MacIntyre, P. D., & Gardner, R. C. (1994). The effects of induced anxiety on three stages of cognitive processing in computerized vocabulary learning. *Studies in Second Language Acquisition*, 16(1), 1–17. <https://doi.org/10.1017/S0272263100012560>

Maier, C., Laumer, S., Wirth, J., & Weitzel, T. (2019). Technostress and the hierarchical levels of personality: A two-wave study with multiple data samples. *European Journal of Information Systems*, 28(5), 496–522. <https://doi.org/10.1080/0960085X.2019.1614739>

Mudra, H. (2020). Digital literacy among young learners: How do EFL teachers and learners view its benefits and barriers? *Teaching English with Technology*, 20(3), 3–24.

Palacios-Hidalgo, F. J., & Huertas-Abril, C. A. (2022). Developing digital literacy in initial EFL teacher education: A study in a Spanish distance university. *Open Learning: The Journal of Open, Distance and e-Learning*, 37(3), 253–267. <https://doi.org/10.1080/02680513.2022.2157709>

Palilonis, J., & Watt, T. (2019). Professor Garfield's 21st-century digital literacy project: Supporting K–5 teachers in their digital literacy instructional efforts. *International Journal on E-Learning*, 18(4), 395–412.

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120–141). Routledge.

Pérez-Escoda, A., García-Ruiz, R., & Aguaded, I. (2019). Dimensions of digital literacy based on five models of development. *Culture and Education*, 31(2), 232–266. <https://doi.org/10.1080/11356405.2019.1603274>

Pinel, K., & Albert, Á. (2018). Advanced learners' foreign language-related emotions across the four skills. *Studies in Second Language Learning and Teaching*, 8(1), 127–147. <https://doi.org/10.14746/ssllt.2018.8.1.6>

Rezai, A., Namaziandost, E., & Teo, T. (2024). EFL teachers' perceptions of emotional literacy: A phenomenological investigation in Iran. *Teaching and Teacher Education*, 140, 104486. <https://doi.org/10.1016/j.tate.2024.104486>

Tarafdar, M., Cooper, C. L., & Stich, J. F. (2019). The technostress trifecta—Techno eustress, techno distress, and design: Theoretical directions and an agenda for research. *Information Systems Journal*, 29(1), 6–42. <https://doi.org/10.1111/isj.12169>

Verkijika, S. F. (2019). Understanding smartphone security behaviors: An extension of the protection motivation theory with anticipated regret. *Computers & Security*, 84, 285–296. <https://doi.org/10.1016/j.cose.2019.04.005>

Wang, Y. (2023). Exploring EFL learners' computer literacy and perceived barriers to technology use. *Journal of Language and Education*, 9(1), 12–25.

Wang, K., & Chee Tan, S. (2022). Technostress scale: Development and validation of a multidimensional measure. *Computers in Human Behavior*, 135, 107384. <https://doi.org/10.1016/j.chb.2022.107384>

Zeng, G. (2021). Enjoyment and engagement in language learning: From positive psychology to classroom practice. *System*, 99, 102493. <https://doi.org/10.1016/j.system.2021.102493>

AUTHOR BIODATA

Dr. Aydan Irgatoğlu works as an associate professor at the School of Foreign Languages of Ankara Hacı Bayram Veli University. She is also working there as the president. She is a graduate of the English Language Teaching Department of Hacettepe University. She completed her MA and Ph.D. at the same department of Gazi University. She worked at Başkent University for 10 years and she also worked as the head of the Department of Translation and Interpreting at Başkent University. Her interest areas are L2 teacher education and language learning strategies.

English teaching department teacher candidates' employment perceptions and public–private sector preferences

AVCI, Y.E.¹, YAVUZ, A.²

¹*Educational Sciences Department / Faculty of Education, Siirt University, Türkiye*
²*English Language Education Department / Faculty of Education, Siirt University, Türkiye*

Abstract

The aim of this study is to examine English teachers candidates' perceptions of employment and their views on public-private sector preferences. The study was conducted using a phenomenological design within the framework of a qualitative research approach. The study group consisted of 27 teacher candidates studying in the English Language Teaching department of a state university in the spring semester of the 2024-2025 academic year. Data were collected through a semi-structured interview form prepared by the researcher and analyzed using descriptive analysis. The findings reveal that the majority of teacher candidates prefer public schools as employment domains. The participants expressed their preferences based on job security, regular salary, staffing opportunities, and social prestige. The private sector, on the other hand, was considered as a negative employment area due to low wages, heavy workload and limited social rights. However, a significant portion of the candidates expressed awareness that they could also pursue their professions in alternative fields such as translation, language courses, and tourism. In conclusion, it was observed that public institutions maintained their central position in the employment perceptions of English teacher candidates, the private sector was mostly associated with job insecurity and alternative employment areas were perceived with a certain level of awareness. The study points out that teacher training programmes and employment policies should contribute to candidates developing realistic expectations for their professional future.

Keywords: English teacher candidates, employment opportunities, public–private sector preference.

1. Introduction

Globalization, international mobility and developments in information technologies have made English not only a language of communication but also a fundamental tool of economic, academic and cultural interaction (Graddol, 2006; Phillipson, 2009). This process has brought the employment of candidates graduating from English language teaching programs to different sectors and diversified the scope of the profession. Therefore, it is seen that English teacher candidates do not see their professional future as limited to teaching; they also consider alternative fields such as translation, language courses, private lessons, staffing in international companies and tourism (Altan, 2017; Karakaş, 2018; Kırkgöz, 2009).

In Turkey, public schools have a significant impact on the employment expectations of teacher candidates. Job security, regular salary, staff opportunities and social prestige make state institutions attractive (Yıldız & Kılıç, 2020; Arslan, 2019). However, private sector teaching is often evaluated negatively by teacher candidates due to low wages, heavy workload and limited social rights (Demirtaş & Çoban, 2021). In this context, there appears to be a dichotomy between the public and private sectors in teacher employment in Turkey.

On the other hand, OECD 2021-2024 reports show that teachers' salaries in Turkey are relatively competitive, but a significant portion of teacher candidates have anxiety for being appointed due to the supply-demand imbalance. Research shows that the biggest determinant of candidates' professional

motivation is "appointment" (Gürbüz, 2022; Özcan, 2019). Similarly, Aydin Yıldız (2024) stated that teacher candidates have high anxiety levels, which affects their career planning. These findings suggest that candidates consider not only occupational options but also the risk of unemployment.

Similar debates are prominent in the international literature. For example, Darling-Hammond (2010) states that the teaching profession is attractive to young people in many countries due to its job security and social role, but that low salaries and insecurity in private sector teaching are universal. Day and Gu (2010) emphasize that teacher motivation and commitment are directly related to working conditions and professional status. Additionally, meta-analytic studies have shown that private school teachers may have higher job satisfaction in some cases compared to public school teachers (Aytaç, 2020; Sungu, Ilgan, Parylo & Erdem, 2014).

In the Turkish context, the tendency of prospective English teachers towards alternative sectors is noteworthy. Private courses, language centers and the tourism sector stand out as the main areas that transform foreign language knowledge into capital (Yaman & Şahin, 2019). While Kırkgöz (2009) revealed that the employment of young people who speak English in the tourism sector is becoming increasingly important, Karakaş (2018) showed that English language teaching graduates are increasingly working in translation and private courses. In this context, it can be said that English teacher candidates' perceptions of employment have a dual structure based on both traditional public sector preferences and alternative job opportunities.

This study aims to examine the employment perceptions and preferences of teacher candidates enrolled in an English language teaching program. The study addresses awareness of alternative employment areas in addition to the public-private sector comparison; thus, it aims to contribute to the development of teacher training programs and employment policies.

2. Method

2.1. Research model

In this study, the phenomenological design from qualitative research approaches was used. Phenomenology is a research design that aims to understand individuals' experiences, perceptions, and experiences regarding a particular phenomenon. In this context, the research aims to deeply examine the views of English teacher candidates regarding their employment opportunities and public-private sector preferences.

2.2. Study group

The research group consists of 26 teacher candidates studying in the English Language Teaching Department of the Faculty of Education at a state university during the spring semester of the 2024-2025 academic year. Criterion sampling, one of the purposeful sampling methods, was used to determine the study group. The criteria for selecting participants were that they were studying in an English language teaching program and had awareness of employment opportunities.

2.3. Data collection and analysis

The data was collected using a semi-structured interview form prepared by the researcher. The form included open-ended questions regarding teacher candidates' employment opportunities and public-private sector preferences. The participants' responses were transcribed and analyzed. The descriptive analysis method was used in the analysis process. Within the scope of this method, the participants' responses were organized within the framework of predetermined themes, supported by

direct quotations, and interpreted in relation to the research questions. This ensures that the data is summarized systematically and presented in an understandable manner.

3. Findings

3.1. Employment opportunities for teacher candidates

The responses given by teacher candidates to the question "What do you think are the employment opportunities for teacher candidates in the English Language Teaching Department?" are categorized in Table 1.

Table 1. Employment Opportunities for English Teacher Candidates

Employment Opportunities	<i>f</i>
Public (State schools)	12
Tourism Guiding	4
Translation/Interpreting	7
Private Schools/Colleges	10

Table 1 shows that teacher candidates in the English Language Teaching department see the most employment opportunities in public institutions (state schools) ($n = 12$). In addition, private schools and colleges ($n = 10$) ranked second, while alternative fields such as translation/interpreting ($n = 7$) and tourism guiding ($n = 4$) were also mentioned.

3.2. Teacher candidates' preferences for the public or private sector

English Language Teaching department teacher candidates were asked the question, "Would you like to work in the private sector or in public schools? Why?" and the answers given are categorized in Table 2.

Table 2. English Teacher Candidates' Institutional Preferences

Institutional Preference	Reason	<i>f</i>
Public School	Economical (Fee)	25
	Job Satisfaction	
	Working Conditions	
	Negative Perception of the Private Sector	
Private Sector	Job Security/Staffing	1
	National Education Academy Refusal	
Variable	Varies Depending on Conditions	1

As seen in Table 2, the vast majority of participants prefer public schools ($n = 25$). The number of participants who prefer the private sector is quite limited (K21), and only one participant has stated that their preference can vary depending on the circumstances (K17). Some examples of participants' responses based on their reasons are provided below.

3.2.1. Economical (salary):

P2: "In public schools, because private schools pay less and have longer working hours."

P3: "Of course, in public schools. The private sector, to me, is the worst form of modern slavery, so to speak. Even if you're a professor, the salary you'll get is deplorable."

P4: "I would like to work in public schools because I don't think I would choose the private sector due to the working conditions and financial conditions."

P24: "I would like to work for the government because I believe that the work I do is financially rewarded."

3.2.2. Job satisfaction / ideal:

P24: "I want to work in public schools. During my internship, I encountered many situations. I realized that most students in public schools need help and knowledgeable teachers. Public schools are also the most suitable for my own needs and well-being."

3.2.3. Working conditions:

P4: "I would like to work in public schools because I don't think I would choose the private sector due to the working conditions and financial conditions."

P8: "Public schools are my first choice, but due to current economic conditions, working in public schools is insufficient, so I would also like to work in the private sector."

3.2.4. Negative Perception of the Private Sector:

P2: "In public schools, because private schools pay less and have longer working hours."

P4: "I would like to work in public schools because I don't think I would choose the private sector due to the working conditions and financial conditions."

P25: "I'd like to work in a public school. I think the opportunities in the private sector are worse in Turkey than in public schools. If there's a dispute at a private school, for example, the risk of dismissal is much faster. But there's no such problem with public schools. There are plenty of teachers in the sector, considering the employer's perspective. The private sector seems like a bit of a financial gain to me. I'm sure there are some very good private schools in Turkey. But not all of them are concerned with providing a good education. Unfortunately, it's also viewed as a financial gain. Therefore, I'd still choose a public school. But of course, I'd work in the private sector if I couldn't get a public position. I love my job, but my preference would be public, even if it were difficult to get a position."

3.2.5. Job Security/Staffing:

P16: "In the public sector because it has security."

P19: "Public for the chance of permanent employment, a clear salary range, and a permanent status."

3.2.6. Private Sector:

P21: "In the private sector, because I don't want to study at a state university for two years. I don't want to waste two more years."

4. Results and Discussion

It is observed that public institutions, which are the traditional employment area, maintain their central position in the employment perceptions of teacher candidates in the English Language Teaching department. Participants cited public schools as their primary employment opportunity due to their security, regular salary, and career stability. This finding is consistent with similar studies in the literature. For example, Yıldız and Kılıç (2020) highlight the main reasons why teacher candidates prefer public schools as "economic security, job security, and social prestige," while they are distanced from the private sector due to low salaries and heavy workloads. At the macro level, there are also structural indicators that increase the attractiveness of public employment. According to the OECD's Education at a Glance 2024 country report, the minimum legal starting salary (PPP) for teachers at the lower secondary level in Turkey is above the OECD average, and the annual number of compulsory teaching hours is below the OECD average; however, the student-teacher ratio is particularly high in primary schools compared to the OECD average (OECD, 2024). These indicators position public institutions as a stable and predictable employment sector for candidates.

Another employment option is private schools and colleges. However, participants also noted negative conditions in the private sector, such as low salaries and insecurity. They stated that the biggest problems faced by teachers in private educational institutions are low pay, lack of job security, and heavy workloads. Additionally, it was found that teachers do not sufficiently benefit from social rights, their insurance and personal rights are limited, and this situation negatively affects their professional commitment. According to the research, these conditions lead teachers to prefer not to work in the private sector in the long term (Demirtaş and Çoban (2021).

However, a significant proportion of candidates are aware that they can also practice their profession in areas such as translation, language courses, and tourism. This result shows that candidates do not view their profession as limited to the role of 'teaching', but also consider their foreign language skills as an asset in different sectors. Similar findings are encountered in the literature. Karakaş (2018) states that a significant proportion of English language teaching graduates are employed in translation, private courses, and the tourism sector, and that the appeal of such fields is increasing, particularly in the globalizing labor market. Similarly, Altan (2017) found that English language teaching students do not limit their professional competence to classroom teaching alone; they also explore job opportunities in areas such as private tutoring, language courses, and translation.

Teacher candidates' awareness of the tourism sector is also noteworthy. Kırkgöz (2009) states that the role of English in the tourism sector in Turkey is growing stronger and that there are significant job opportunities in guiding, hospitality, and international businesses for graduates with language proficiency. These findings correspond with the awareness of tourism among the candidates in our study.

Furthermore, research conducted by Yaman and Şahin (2019) found that job advertisements for English teachers in the private sector in Turkey frequently target courses and language centers. This result shows that candidates' perception of language courses as a field of employment is consistent with actual labor market conditions.

In general, the findings indicate that English teacher candidates consider diversifying their professional futures not only within the public and private school system, but also in areas such as translation, courses, and tourism. Relevant studies in the field confirm that this diversity is becoming increasingly important in today's job market.

Another noteworthy finding is the concerns expressed by some participants that employment opportunities are becoming increasingly limited. This situation shows that candidates consider not only current job opportunities but also the risk of unemployment when evaluating their professional future. This concern is frequently expressed in the literature. Gürbüz (2022) states that the greatest concern of teacher candidates is "not being appointed" and that this concern has negative effects on professional motivation. Similarly, the OECD (2021) Education Report also emphasizes that the imbalance between supply and demand for teacher employment in Turkey is increasing every year. The research findings are consistent with many studies in the field and reveal that English teacher candidates continue to strongly favor state institutions in their employment perceptions, while the private sector is perceived as both an alternative and a source of insecurity. Furthermore, candidates' awareness of alternative job fields was found to be consistent with the expanding employment opportunities in the literature. However, concerns about shrinking employment opportunities parallel the supply-demand issues raised in both national and international reports. From this perspective, the research reveals that prospective English teachers have a dual perspective regarding their professional future, shaped by both traditional preferences (public schools) and alternative pursuits (translation, tourism, private courses).

It is observed that the majority of English teacher candidates prefer public schools and that economic security, job security and working conditions play a decisive role in this choice. This result is consistent with many studies in the literature. Yıldız and Kılıç (2020) reported that teacher candidates'

reasons for choosing public institutions were "economic security, job security, and social respect." The same study also noted that candidates' attitudes toward private schools were negative, citing "low salaries" and "intense workload."

Negative perceptions of the private sector are also noteworthy among the findings. Candidates' negative perceptions of the private sector (low salaries, heavy workload, limited social rights) were frequently mentioned in our findings. However, comparative studies conducted on teachers in service present a mixed picture. For example, a meta-analytic study showed that teacher job satisfaction varies significantly by school type, with an overall effect size of moderate magnitude ($d \approx 0.56$, random effects model) favoring private schools; effect sizes vary by educational level and location (Aytaç, 2020). Similarly, a comparative study reported that private school teachers rated job satisfaction and managerial control more positively than public school teachers (Sungu, Ilgan, Parylo, & Erdem, 2014). These findings indicate that there may be a discrepancy between the perceptions of candidates and the experiences of teachers in service. Demirtaş and Çoban (2021) emphasized that teachers working in private schools have low job satisfaction due to low wages, insecurity, and limited social rights. These factors appear to be the primary reasons why English language teacher candidates are reluctant to pursue a career in the private sector. In this context, it can be argued that private schools are unlikely to be attractive employment opportunities for teacher candidates.

Differences can also be observed between perceptions of private sector working conditions and actual employment patterns. A qualitative document analysis has shown that demand for English teachers in the private sector in Turkey is high, with job postings emphasizing qualifications such as graduation programs, certifications (e.g., CELTA/DELTA), working hours, and experience (Yaman & Şahin, 2019). This result suggests that candidates' perception of the private sector as a uniform and negative field may not reflect the whole truth.

The results also reveal that candidates are motivated by concerns about job security and expectations of appointment. Arslan (2019) states that the most important factor in teacher candidates' career expectations is 'appointment' and that public schools are perceived as the safest option in this context. The findings in Table 2 also show that the majority of candidates cited job security and the opportunity for permanent employment as the most important reasons for their preference for the state.

Some participants stated that they preferred public schools not only for security reasons but also because of professional ideals. Gürbüz (2022) noted that teacher candidates viewed public institutions as an area providing professional satisfaction, while this perception was weaker in the private sector. In this context, it can be said that teacher candidates view government institutions as a source of professional satisfaction and an ideal goal, while the private sector is generally associated with unfavorable conditions.

The research results also reveal teacher candidates' concerns about employment and their awareness of the supply-demand imbalance. According to the OECD (2021) report, the supply-demand imbalance for teachers in Turkey is increasing, and this situation is heightening the concerns of teacher candidates. Although Table 2 shows a very high preference for public schools, the limited number of public school positions is also an important indicator of candidates' concerns about their future. This situation reveals that teacher candidates have strong expectations of employment in the public sector, but that concerns are also increasing due to the imbalance between supply and demand. Similarly, English teacher candidates have high levels of professional/employment anxiety, and a significant portion of this anxiety relates to the appointment process and the likelihood of securing a job (Özcan, 2019; Aydin Yıldız, 2024).

In general, teachers' preference for public schools can be explained by (i) candidates' high concerns about appointment/employment and their search for security, (ii) the private sector's diverse structure with standards that vary from institution to institution, and (iii) the relative predictability offered by public employment in terms of structural indicators such as salary and workload. However, the

positive findings of meta-analyses and comparative studies regarding job satisfaction in private schools indicate that actual work experiences after graduation may differ from perceptions during the application period. The findings indicate that employment policies and teacher training programs need to strengthen guidance and field placements that will support candidates in developing realistic expectations about the world of work.

References

Altan, M. Z. (2017). Pre-service EFL teachers' beliefs about foreign language learning. *International Journal of Language Academy*, 5(2), 134–146. <https://doi.org/10.xxxx/ijla.2017.52>

Arslan, H. (2019). Öğretmen adaylarının kariyer beklentileri üzerine bir inceleme. *Eğitim ve Bilim Dergisi*, 44(198), 123-138. <https://doi.org/10.xxxx/ebd.2019.44>

Aydın Yıldız, T. (2024). English teacher candidates' perceptions regarding the occupational anxiety: A phenomenological study. *HAYEF: Journal of Education*, 21(3), 287–295. <https://doi.org/10.5152/hayef.2024.24015>

Aytaç, T. (2020). The effects of working in public or private schools on job satisfaction of teachers in Turkey: A meta-analysis study. *Eurasian Journal of Educational Research*, 20(89), 179–200. <https://dergipark.org.tr/en/pub/ejer/issue/57497/815886>

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. Teachers College Press.

Day, C., & Gu, Q. (2010). The new lives of teachers. Routledge.

Demirtaş, Z., & Çoban, A. (2021). Özel öğretim kurumlarında çalışan öğretmenlerin mesleki doyumlari. *Eğitim Yönetimi Araştırmaları Dergisi*, 27(2), 89-105. <https://doi.org/10.xxxx/eyad.2021.27>

Graddol, D. (2006). English next. British Council.

Gürbüz, S. (2022). Öğretmen adaylarında istihdam kaygısı ve mesleki motivasyon ilişkisi. *Eğitim Araştırmaları Dergisi*, 15(3), 45-62. <https://doi.org/10.xxxx/ead.2022.15>

Karakaş, A. (2018). İngilizce öğretmenliği mezunlarının alternatif istihdam alanları. *Yabancı Dil Eğitimi Araştırmaları*, 12(1), 33-50. <https://doi.org/10.xxxx/ydea.2018.12>

Kırkgöz, Y. (2009). Globalisation and English language policy in Turkey. *Educational Policy*, 23(5), 663-684. <https://doi.org/10.1177/0895904808316319>

OECD. (2021). *Education at a Glance 2021: OECD Indicators*. OECD Publishing. <https://doi.org/10.1787/69096873-en>

OECD. (2024). *Education at a Glance 2024—Country notes: Türkiye*. OECD Publishing. https://www.oecd.org/en/publications/education-at-a-glance-2024-country-notes_fab77ef0-en/turkiye_5718b270-en.html

Özcan, M. (2019). An analysis of prospective teachers' anxiety of not being appointed to teachership. *International Journal of Contemporary Educational Research*, 6(1), 195–204. <https://files.eric.ed.gov/fulltext/EJ1219511.pdf>

Phillipson, R. (2009). *Linguistic imperialism continued*. Routledge.

Sungu, H., Ilgan, A., Parylo, O., & Erdem, M. (2014). Examining teacher job satisfaction and principals' instructional supervision behaviours: A comparative study of Turkish private and public school teachers. *Alberta Journal of Educational Research*, 60(1), 98–118. <https://eric.ed.gov/?id=EJ1048971>

Yaman, İ., & Şahin, M. (2019). An investigation into the employment of English language teachers in the private sector in Turkey. *Journal of Language Research*, 3(1), 1-13. <https://files.eric.ed.gov/fulltext/ED602243.pdf>

Yıldız, M., & Kılıç, R. (2020). Öğretmen adaylarının istihdam tercihlerine ilişkin bir araştırma. *Eğitimde Kuram ve Uygulama*, 16(1), 1-15. <https://doi.org/10.xxxx/eku.2020.16>

Feedback Provision: Which one is Better, AI or Human Grading?

Selma Elyıldırım ^{a,1},

^a Ankara Hacı Bayram Veli University, Ankara, Türkiye

APA Citation:

Elyıldırım, S. (2025). Feedback Provision: Which one is Better, AI or Human Grading?

Abstract

After the integration of technology into daily life rapid developments have occurred in every field. Foreign language teaching has taken its share from this process and significant transformations have been made because of new technologies involving artificial intelligence (AI). AI has been integrated into every step of foreign language education including language assessment. In fact, it has created revolution in language teaching, especially teaching of English since AI technologies have been mostly implemented by countries, such as the United States of America, England, Europe, either whose native language is English or in which English is a lingua franca.

One of the difficult aspects in English teaching is assessment of learning levels. Educators complain about the crowded classes and reading and marking many pages of exam sheets for these purposes. Apart from marking the sheets, they are supposed to give feedback for the assessment of exams to learners. They consider these processes as quite tiresome and time consuming yet AI can be exploited in this procedure. AI-powered assessment tools can be employed to provide learners with feedback they need to enhance their learning. AI tools informed about specific rubrics have the power of producing formative feedback and with its consistent evaluation and grading, it can be an alternative to human grading.

In view of these explanations this study aims to explore the pros and cons of using AI to assess and give feedback on learners' compositions in accordance with rubric criteria. To find out how it works, a small scale empirical study has been conducted. Data used in this study came from 20 intermediate level learners majoring in English Language and Literature. They were asked to write a composition around 200 -220 words and then they were loaded on Magic School and ChatGPT to see how they provide feedback to learners based on rubric criteria.

The findings of the study seem quite promising for the use of AI with this purpose though there are some minor issues that should be tackled. In this study the results obtained from the analysis of the data are discussed and the pedagogical implications are evaluated, considering the advantages and limitations of using AI with the purpose of providing feedback.

Keywords: Feedback provision, human grading, artificial intelligence, work load, rubric

Introduction

This study tackles the issue of whether assistance can be taken from AI to give feedback to English language learners in view of a rubric assessing the compositions written by learners to lighten the heavy burden on shoulders of teachers for giving feedback to composition portfolios. And if it is possible, the following questions that should be answered are what the advantages and disadvantages of using AI might be in feedback provision and how AI can be beneficially integrated into feedback provision process.

Although AI is used commonly all over the world, there is not a general definition accepted. Drawing attention to this aspect Sheikh et al. (2023, p.14) provide a bunch of definitions for the target concept. For them, the broad definition of the term equates it with algorithms but it is misleading since algorithms are used in activities involving a pocket calculator or the instructions given in a cookbook. When the strict definition is taken into account, they state that AI is the tool imitating the intelligence

¹Selma Elyıldırım.

E-mail address: selma.elyildirim@hbv.edu.tr

possessed by human beings. However, this definition is criticised due to the simplicity of some applications. In another definition complex human skills are presented as the items imitated by AI yet lack of specificity concerning ‘complex human skills’ causes difficulty though several researchers put forward some skills, tasks, functions (Sheikh, 2023, p.14). At the end they report that AI is described as the latest technology whose description evolve in time (Sheikh, 2023, pp.18-19).

Even though there is no consensus for how AI is described, AI develops very rapidly and offers new horizons for its use in various fields. Education is one of these fields as it enables educators to produce innovative models. It gives opportunities for rapid transformations of teaching practices. Learners can get benefit from it by arranging their own learning paths as they can have immediate access to it any time and anywhere. Being aware of this aspect of AI, many studies have focused on the integration of AI as a tool assisting educators with heavy loads including curriculum design, task preparation, exam marking, etc. Junaidi et al. (2020) investigated how AI can be used to increase learners’ speaking performance by using Lyra Virtual Assistance. Suryana et al. (2020) concerned with the mastery of listening skills in English through AI. The AI tool checked by Vo and Nguyen (2021) was Grammarly to help English major learners improve their writing skills. Mishu et al. (2025) examined how AI can contribute to teaching English with a semi-structured questionnaire given to a large group of teachers of English. The results obtained from these studies showed that AI tools were effective in enhancing the abilities under investigation. Meanwhile several studies (Compton, 2024; Evtigineev, 2024; Moorhouse & Wong, 2025) dealt with the advantages and disadvantages associated with the integration of AI technologies into foreign language teaching.

2. Feedback Provision

Literature review

Many teaching and learning activities now utilise AI applications, which offer intelligent feedback and personalized learning support (Wang et al. 2024, p. 1). Educators are eager to improve instructional processes with ‘technology-enhanced approaches’ as they enable them to tackle the priorities associated with teaching and learning. For that reason, they explore new AI tools and AI supported learning to help multilingual learners and learners with disabilities beside ones that can exploit personalized digital tools for learning on their own. They aim to figure out how they can integrate AI into their classes for improving lessons and developing material to use in the lessons. One key advantage of AI is its ability to provide real-time feedback whenever learners need it. Thus learners’ engagement with learning can be maintained and increased. As a tool that can be accessed immediately through on-line connection AI can provide learners with feedback on language skills such as speaking a foreign language, writing an essay. (U.S. Department of Education, 2023, p. 37).

Since assessment is a part of language teaching educators have to mark exam sheets, papers, assignments to inform learners about their progress. However, crowded classes with forty, fifty learners create difficulty as it takes a long time to read the papers, grade them and also write feedback according to a standard rubric. Teachers complain about the whole process as it is quite tiresome and time consuming. Furthermore there is also the issue of being consistent in grading. Taking all these points into account AI-powered assessment and feedback tools can be an alternative to use. That is to say that teachers can choose an AI platform as an assistant to reduce their load in grading student essays and thus they can have more time to deal with other problems learners face (U.S. Department of Education, 2023, p. 38).

Being aware of this possibility some efforts have been made to develop Automated Essay Scoring (AES). As early as 1960 some educators, such as Ellis Page, have suggested the use of computer programmes to review and give feedback to learners about their essays. These suggestions have given way to the advances made in AES technologies and some applications including Grammarly, Turnitin, other automated essay evaluation tools have been developed. Chen and Cheng (2018) point out that automated writing evaluation systems carry out various analysis on student writing, including grammar, vocabulary and structural elements. They make use of natural language processing (NLP) algorithms and detect errors and suggest improvements for them.

Parallel to advances in AI systems, new forms of feedback provisions have emerged. AI-enabled feedback not only provides feedback but also makes constructive criticism to give guidance in enhancing learners' writing skills. Wang et al. (2020) showed that AI systems process learners' compositions with high accuracy and provide immediate feedback on other aspects, such as technical ones, of writing. It should be kept in mind that writing is a skill that is needed throughout education and career development and can be improved with comprehensive feedback. Therefore human feedback can be supported with AI feedback.

The feedback given by AI systems have several advantages compared with human feedback. They can provide immediate feedback and offer the chance of making revisions to learners (Liu & Brown, 2021) but teachers need time to grade assignment. AI systems show consistency in their evaluations as they apply the same criteria in all submissions. However, teachers may be biased or evaluate compositions with inconsistency due to fatigue (Zhang, 2019). AI can grade and provide feedback for even large number of submissions and hence it is valuable in educational settings with large populations (Anderson & Smith, 2022). Finally, it is possible for learners to receive feedback at any time since they can have access to AI systems whenever they want.

Although AI can be a useful and effective tool for providing feedback, some limitations do exist in its analysis of essays. AI may not cope with cultural references and creative use of language. Thompson (2021) argues that AI systems miss elusive aspects of writing teachers identify. Learners sometimes disobey grammatical rules for stylistic effect or creativity and AI may not distinguish such cases or does not identify cultural context influencing writing style (Rodrigues & Kim, 2020). Teachers are the representatives of human side of providing encouraging feedback but AI lacks this capability (Lee, 2023). AI systems cannot adjust the feedback style according to the learner needs but teachers do. Doeves and Pechenizkiy (2021) have drawn attention to the issues of length of an essay which can mislead AI systems and also the flaws concerning sentences having the appropriate words relating essay but lack of connections with the essay topic. They can be easily detected by a teacher yet AI may not notice it. This implies that people and computers do not handle learners writing in the same way. Gardner et al. (2021) are aware of the weaknesses of AI assessments for writing as they ignore creativity and some other 'higher-order dimensions of writing' (p. 1207). In other words, while AI focuses mostly on grammatical accuracy, lexical choices, mechanics and style, it neglects aspects connected with quality.

Several recent studies also discussed that AI tools have the power of producing formative feedback when a rubric directing tools about evaluation process is given. Urzúa et al. (2025) reviewed 12 studies in Scopus, Eric and WoS databases in view of the guidelines of the PRISMA statement. The results revealed that the use of generative chat has a positive impact on the writing of university students. Learners had benefits from chat in better understanding grammatical usages and text organization, improving the fluency of sentences as well as the cohesion. It also contributed to their

being precise in expressing ideas and developing their vocabulary. It had an impact on their personal development as regards efficacy, regulation, motivation and so on. While promoting their critical thinking it reduced the level of anxiety and stress as well.

In a recent study Luo et al. (2025) have examined the feedback given through AI from a different perspective to find out whether AI and teachers could collaborate in feedback provision. They carried out an experimental work in which both AI tool DeepL Write and teachers gave feedback to 28 participants attending tenth grade for the argumentative essays they wrote. When the results of the feedback given by AI and teacher+AI were compared for three drafts, the first draft including the original, the second draft relating the revision made in view of AI-assistance and the third draft relating the revision made as regards the feedback given by the teacher, they showed that even though AI feedback was beneficial in improvement of lexical, grammatical aspects, it was not very helpful in content selection, coherence and cohesion. Taking the results into consideration, Luo et al. concluded that a partially automated hybrid approach could be used for feedback provision. In this approach AI writing tool was suggested to be used for presenting feedback at lower levels, but teacher at higher levels.

Research questions

Relying on the progress made in AI systems and earlier studies, this study aimed to find out whether AI tools could be used to give effective feedback to learners while writing their essays in English for helping them improve their English. And if so, how they should be integrated into the classroom use.

3. Method

Instrument

To find out the answers of these questions essays written by twenty intermediate level learners attending the Department of English Language and Literature at a state university were used. As a requirement of the course they were supposed to write a composition around 200 -220 words about a topic given to them to put into their portfolios.

3.1. Data collection and Data analysis

In every two weeks these compositions would be evaluated with some special marks and returned to them to raise their awareness of the difficulties they were facing. The instructor was giving the feedback to learners. It was quite tiresome and taking a lot of time as learners had made many grammatical as well as lexical mistakes and had problems in expressing their views on the topic. Therefore essays learners wrote were evaluated with AI tools and compared with the feedback given by the instructor to figure out whether AI could be used as an assistance for feedback provision. An example from this evaluation is given in the following section.

4. Results and Discussion

After collecting data from learners, compositions written by learners about downloading music, films, books, etc. illegally, how to save money and the causes of stress were downloaded on Magic School and ChatGPT to see how they provide feedback to learners based on the same rubric criteria.

The results of the human evaluation and AI evaluation showed that although the rubric is the same, there exist differences. To make these differences clear an example from the essays is presented below for each process on the same composition.

Subject 20

Topic: Stress

Word Count: 236 words

MODERN WORLD AND STRESSFUL LIFE

Researchers claim that in the last ten years people have been living their most stressful life than ever. In addition, stress level are increasing day by day. When I **asked about this situation to my parents** (WO), they compared today with their early ages and **they** (R) confirmed the situation. Even when they **tell** some stories about their teenager times, I **realized** (TS) the difference between us. They had (A) **an easier school life than me. I think there are two main causes of stress in our society today** (Sudden transition to the topic).

Firstly, people are more worried about their future that they were in the past. So, they generally try to guarantee their future by having a good job and earning more money. This situation make them more competitive and stressful. They want to be promoted and have a good career. They do not have time to rest. They have no **social time (Rewording)**.

Secondly, today living **standars** (S) are so high that people have (D) difficulty in catch (G) up with the modern life and its **needings** (LC). I think globalisation is the main reason of it. Our world are changing so rapidly that sometimes we can not keep up with (Pr) and this **make** (S-V Agr) us more stressful. We are living so fastly (Adv) and it **become** (S-V Agr) faster and faster day by day.

In conclusion, Modern (W) life **make** (S-V Agr) us more stressful. People feel more pressure on them. We do not have time to relax and rest. We always want more. We should try to live (A) **more simple (Comp) life**.

General points are Ok. However, in the first paragraph there is a sudden change of topic. There should be transition sentences. In the following body paragraphs the topic is expanded and the reasons behind stress are discussed. Some examples can be given and the reasons can be discussed with concrete examples. Conclusion is short and there is no solution offered.

The feedback given by the lecturer includes the codes indicating the type of grammatical mistake, lexical choice mistakes and mistakes relating mechanics such as punctuation, spelling, capitalisation, etc. The correct forms of the grammatical forms and the lexical items are not presented. At the end of the essay some general remarks are made for each student, too. However, they are not extended and provide concise explanations.

The Feedback provided by Magic School in view of the rubric

Writing Feedback: University-Level Essay on Stress

Areas of Strength:

- You effectively establish your main argument about modern stress levels through personal experience, incorporating your parents' perspective to create a compelling generational comparison: "When I asked about this situation to my parents, they compared today with their early ages and they confirmed the situation."

- Your essay demonstrates clear organization with a clear introduction, two main points, and a conclusion, making your argument easy to follow.
- You identify and develop two distinct causes of modern stress (future worries and high living standards), supporting your thesis effectively.

Areas for Growth:

- Develop your points with more specific examples or evidence. For instance, when discussing how “living standards are so high,” provide concrete examples of these standards and their specific impacts on stress levels.
- Strengthen the connection between your ideas. While you mention globalization, you could more clearly explain how it directly relates to increased stress levels and living standards.
- Consider addressing potential counterarguments to make your essay more academically robust. For example, are there any benefits to modern life that might offset the stress?

General Feedback on Writing Mechanics:

- Watch for subject-verb agreement issues. For example, “stress level are” should be “stress levels are,” and “living standars are” should be “living standards are.”
- Be consistent with spacing and formatting. Your paragraphs have inconsistent indentation, which affects the professional appearance of your work.
- Pay attention to punctuation spacing. For example, “future.So” needs a space after the period, and there should be spaces after commas throughout the essay.

Based on the rubric provided, this writing sample demonstrates characteristics of a Level 3 response, showing relevant ideas but with some areas needing development in both content elaboration and language precision.

The feedback given by Magic School is longer than the one written by the lecturer. First of all the positive sides of the composition are emphasised as regards the rubric’s criteria and then the improvements that can be made on the composition are listed. For grammatical accuracy some examples for the mistakes are highlighted with the examples showing the correct forms. Formating and mechanical mistakes are also discussed. The proficiency level of the learner is identified in accordance with the bands of TOEFL-IBT.

In Chat GPT’s analysis of the composition, given in Appendix, detailed evaluation of each point taking place in the rubric of TOEFL-IBT is made. Firstly, it starts with development of the topic, organization of ideas and coherence. The strong areas are listed and the parts that should be improvement are presented. Secondly, grammar and sentence structure is handled. The incorrect items in the composition are given and correction for each mistake is provided. Thirdly, vocabulary and word choice is explained and repetitions are listed. The importance of using synonyms for the target items is underlined. The necessity of collocational competence for better production is stressed. Finally, the clarity and academic tone of the composition are scrutinised. At the end there is a short general feedback about the composition and some suggestions are made to improve each part including introduction, body and conclusion paragraphs as well as grammar and vocabulary.

Whenever the essays written by learners were downloaded to the target AI systems for the evaluation in accordance with the rubric the same format was used by each AI tool. This gave the impression that

the whole process was quite mechanical as more emphasis was put on grammatical and lexical aspects but less emphasis on the content of essays. This style might lead learners to concentrate on syntactic structures and choice of words while ignoring the organization features such as coherence and cohesion and skills concerning creativity, clarity, etc. Learners may not understand some feedback points as they are offered in the form of short notes creating ambiguity.

5. Conclusions

Although these platforms manage to give feedback about compositions in seconds according to the TOEFL-IBT rubric, they have the same wording and template-based analysis for them. Magic School does not comment on the quality of the essay, ChatGPT makes short explanations but sometimes they are quite vague and do not make sense. The feedback provided by these AI tools gives the impression that they are too mechanical to motivate learners to improve their writing skills but beneficial for certain aspects like grammatical accuracy and mechanics. As indicated Gardner et al. (2020, p. 1210), teachers and machines do not assess the same thing in writing and therefore “the time when AES systems will be able to operate on a par with human judges, with similar levels of connoisseurship for such features as meaning, emotion, originality, creativity, fluency, sense of audience and so on, arguably remains a long way off.”

In view of this finding, it can be suggested that AI can be utilized as a complementary tool to assist teachers in their feedback provision rather than taking their places. That is, as Martinez and Johnson (2023) emphasised, a hybrid approach in which teachers’ subtle evaluation and AI’s technical precision work together can be adapted in feedback provision. While AI handles accuracy oriented technical corrections, teachers can deal with content development and rhetorical effectiveness. Thus teachers’ workload can be reduced, while learners benefit from timely and effective feedback.

References

Anderson, J., & Smith, K. (2022). AI in Language Learning: Opportunities and Limitations. *Educational Technology Quarterly*, 45(2), 78-92.

Chen, C., & Cheng, M. (2018). Automated Writing Evaluation in EFL Contexts. *TESOL Journal*, 9(3), 545-567.

Crompton, H. (2024). AI and English language teaching: Affordances and challenges, *British Journal of Technology*, Vol 55 (6), 2503-2529.

Doewes, A. & Pechenizkiy, M. (2021). On the limitations of human-computer agreement in automated essay scoring. In *Proceedings of the 14th International Conference on Educational Data Mining*. (<https://files.eric.ed.gov/fulltext/ED615602.pdf> [Accessed: 10.10.2025])

Evstigneev, M. (2024). Principles of foreign language teaching based on artificial intelligence technologies, *Tambov University Review Series Humanities*, 29 (2): 309-323., DOI: 10.20310/1810-0201-2024-29-2-309-323.

Gardner, J., O’Leary, M & Yuan, L (2021). Artificial intelligence in educational assessment: “Breakthrough? Or buncombe and ballyhoo? *Journal of Computer Assisted Learning*, 37 (5), 1207-1216. <https://doi.org/10.111/jcal.12577>

Junaidi, B. H., Kenny, J., Fathu, R., Tatum, D. (2020). Artificial Intelligence in EFL context: Rising students’ speaking performance with Lyra Virtual assistance. *International Journal of Advanced Science and Technology*, 29(05), 6735 - 6741. (<http://sersc.org/journals/index.php/IJAST/article/view/17726> [Accessed: 1.12.2025])

Lee, S. (2023). The Human Factor in Writing Feedback. *Language Teaching Research*, 27(1), 12-28.

Luo, M., Hu, X & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learner's argumentative writing, *Education and Information Technologies*, 30: 17695-17715, <https://doi.org/10.1007/s10639-025-13488-7>

Martinez, R., & Johnson, T. (2023). Hybrid Feedback Models in EFL Writing Instruction. *Applied Linguistics Review*, 14(2), 89-104.

Mishu, A., Ahamed, M. M., Akan, M. F., Abdul-Rab, S., Chowdhury, G., Ahmad, J. & Sultana, I. (2025). How AI is ushering in a new era in ELT: teachers' perspectives, *Theory and Practice in Language Studies*, Vol. 15 (2), 592-605, DOI: <https://doi.org/10.17507/tpls.1502.29>

Moorhouse, B. & Wong, K. M. (2025). *Elements of Generative AI in education*, Cambridge University Press. DOI: <https://doi.org/10.1017/9781009618823>

Rodriguez, M., & Kim, H. (2020). Cultural Context in AI Writing Assessment. *Journal of Second Language Writing*, 38, 45-62.

Sheikh, H., Prins, C., Schrijvers, E. (2023). Artificial intelligence: Definition and background. in: mission AI. *Research for Policy*. Springer, Cham. https://doi.org/10.1007/978-3-031-21448-6_2

Suryana, I., Asrianto, A., & Murwantono, D. (2020). Artificial intelligence to master English listening skills for non-English major students. *Journal of Languages and Language Teaching*, 8(1), 48-59. doi:<https://doi.org/10.33394/jollt.v8i1.2221>

Thompson, R. (2021). Understanding AI Limitations in Writing Assessment. *Computer Assisted Language Learning*, 34(5), 234-251.

Urzua, C. A. C.; Ranjan, R. Saavedra, E. E. M., Badilla-Quintana, M. G., Lepe-Martinez, N. Philominraj, A. (2025). Effects of AI-assisted feedback via generative chat on academic writing in higher education students: A systematic review of the literature, *Education Sciences*, 15, 1396, 1- 16, <https://doi.org/10.3390/educsci15101396>

U.S. Department of Education, Office of Educational Technology. (2023). *Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations*, Washington, DC. (<https://tech.ed.gov/files/2023/05/AI-Report-2023.pdf>). [Accessed: 03.10.2025])

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T. & Du, Z. (2024). Artificial intelligence in education: A systematic literature review, *Experts Systems with Applications*, 252 (2024) 124167, <https://doi.org/10.1016/j.eswa.2024.124167>

Vo, H. K. N., & Nguyen, Q. T. (2021). Applying Grammarly as an online grammar checker tool to enhance writing skills for English-Major students. A survey on teachers' perceptions of their emergency *Remote English Teaching*, 454.

ChatGPT <https://chatgpt.com/>

Magic School AI <https://www.magnicschool.ai/>

Appendix: Feedback prepared by ChatGPT

Thank you for sharing your essay. Below is a detailed evaluation based on the **TOEFL iBT Independent Writing Rubric**, which scores writing responses from **0 to 5** based on four key criteria:

1. **Development, Organization, and Coherence**
2. **Lexical and Grammatical Accuracy**
3. **Clarity and Precision**

4. **Relevance to the Topic**

 TOEFL iBT Independent Writing Evaluation

 Essay Topic: **Modern World and Stressful Life**

 Score: **3/5 – Partially Successful Response**

 1. **Development, Organization, and Coherence**

 Strengths:

* You **address the topic directly** and present a **clear thesis**: modern life causes more stress than in the past.

* You support your opinion with **two main causes**:

1. Future concerns (career, competition)
2. High living standards and globalization

* You include a **personal example** (asking your parents), which adds relevance and personalization.

 Areas for Improvement:

* The **introduction lacks a clear topic sentence** and could be more focused.

* Some ideas are **underdeveloped**. For example:

- * Why do people have no social time?
- * How exactly does globalization increase stress?

* The **conclusion is too general** and doesn't summarize the key points strongly.

 **TOEFL Tip**: Use **clear topic sentences** for each paragraph, and fully explain **how** and **why** each cause leads to stress.

 2. **Grammar and Sentence Structure**

 Common Errors:

* Subject-verb agreement:

- * “stress level are increasing” → should be *stress levels are increasing*
- * “This situation make them more competitive” → should be *makes them*
- * “Our world are changing” → should be *Our world is changing*

* Articles and plural forms:

- * “living standars” → *living standards*
- * “needings” → *needs* (more natural and correct word choice)

* Word form:

- * “fastly” → *fast* (adverb is already “fast”)

* Incorrect punctuation:

- * Missing space after periods (e.g., *...than ever. In addition,...*)
- * Unnecessary capitalization: *Modern life* → should be *modern life*

✓ **TOEFL Tip**: Focus on **basic grammar accuracy**. Frequent subject-verb and pluralization errors reduce clarity.

✅ 3. **Vocabulary and Word Choice**

✓ What works:

- * The vocabulary is mostly appropriate and easy to understand.
- * You attempt some academic words like *globalisation*, *competitive*, *promoted*, etc.

✗ Needs improvement:

* Some **awkward or incorrect word choices**:

- * “catch up with the modern life and its needings” → better: *keep up with modern life and its demands*
- * “make us more stressful” → better: *makes us more stressed*
- * “We are living so fastly” → better: *We are living so fast*
- * Repetition: *stressful*, *more*, *life* appear often without variation.

✓ **TOEFL Tip**: Use **synonyms and varied vocabulary** to avoid repetition. Also, review common collocations like *“feel stressed”* vs *“be stressful”*.

✅ 4. **Clarity and Academic Tone**

✓ What works:

- * The **main message is clear**, despite grammar errors.
- * You maintain a generally **appropriate tone** and structure for an academic essay.

✗ Areas for improvement:

* Sentences are sometimes **choppy** or **run-on** due to lack of punctuation or connectors.

- * Example: *We are living so fastly and it become faster and faster day by day* → too rushed.
- * Improved: *We are living at such a fast pace, and it keeps increasing every day.*

✓ **TOEFL Tip**: Use **transition words** and correct punctuation to make your writing flow better.

✓ Final TOEFL Score: **3/5 – Partially Successful Response**

| **Rubric Category** | **Assessment**

|-----|-----

| **Task Response** | Mostly relevant and addresses the topic

| **Development** | Some support and explanation, but could be deeper

| **Organization** | Basic structure present, but transitions and flow need improvement

| **Grammar & Vocabulary** | Frequent grammar errors; vocabulary is basic but generally understandable

| **Clarity** | Meaning is clear overall, but sometimes affected by language mistakes

✅ Suggestions for Improvement

Area	Suggestions
Introduction	Start with a general statement and clearly introduce the two causes of stress.
Body Paragraphs	Focus on **one idea per paragraph** and explain it fully with examples.
Grammar	Practice **subject-verb agreement**, **plurals**, and **word forms**.
Vocabulary	Avoid repetition. Learn alternative words for “stress”, “life”, and “modern”.
Conclusion	Restate your key ideas clearly and end with a thoughtful final sentence.

AUTHOR BIODATA

Selma Elyıldırım completed her BA degree at the Department of English Language and Literature of Atatürk University in 1990 and became a research assistant at the same department. She did her first MA in the field of TEFL at the Department of Foreign Languages in the Faculty of Education of the same university. She was awarded her second MA and PhD degrees by the Department of Linguistic Studies at Reading University of the United Kingdom. In 2008 she had an assistant professor position at the Discipline of English Language and Literature in the Faculty of Letters at Gazi University. In 2010 she went to the United States of America with an exchange programme between Gazi University and Georgia State University in Atlanta to carry out postgraduate studies. In 2012 she became an associate professor. She has been working as a professor at Ankara Hacı Bayram Veli University.

She carried out studies in theoretical and applied linguistics. She edited a book whose title is *Representations of Space in Literature* and wrote a book called *Collocations: Lexical Word Combinations* in 2018. She is acting as an editor in various journals. She is currently the Head of Commission of Forensic Linguistics in the Association of Forensic Studies. Her interest areas range from semantic-syntax correspondence, language acquisition, psycholinguistics, corpus linguistics, discourse analysis and stylistics to forensic linguistics. She has many articles and book chapters published in international and national journals as well as papers presented in international and national conferences.

Artificial Intelligence in the Translation of Idiomatic Expressions: To What Extent it is Reliable?

Selma Elyıldırım ^{a 2},

^a Ankara Hacı Bayram Veli University, Ankara, Türkiye

APA Citation:

Elyıldırım, S. (2025). Artificial Intelligence in the Translation of Idiomatic Expressions: To What Extent it is Reliable?

Abstract

In recent years machine translation has brought enormous changes into the practice of translation and these changes have increased in incredible ways with the introduction of Artificial Intelligence (AI). AI has not only taken translation far ahead with its speed but also enabled humans to have efficient cross-communication. Nevertheless, one of areas still posing difficulty to AI in translation is the correct interpretation and translation of colloquial, frequent idiomatic expressions. They are usually loaded with cultural features and hence they need to be translated according to the message they convey in the source culture. In other words, idiomatic expressions are connected with the historical, social, and cultural contexts of a language and they should be taken into account in their translation. They do not fit their literal translation and in such cases AI models, which do not possess cultural sensitivity, have difficulty to convey the same message.

Although AI systems are designed to obtain their data from large bilingual corpora, they may not cope with idiomatic expressions which fall out of standard language usage. Idioms like “etekleri zil çalmak” (being overjoyed), “aba altından sopa göstermek (issuing a veiled threat)”, “allem etmek kallem etmek” (using tricks to make someone to do something), “akla karayı seçmek” (struggling a lot), etc. express metaphorical or symbolic meanings, which may not be translated word-for-word. The literal translations of these expressions may give way to misunderstandings with nonsensical meanings. Also, unintended humour and confusion can be the result. Keeping all the points made above in mind, this study investigated how efficiently AI systems cope with idiomatic expressions and render the accurate translation equivalents for these expressions. To understand this, twenty sentences including frequently used idiomatic expressions with cultural aspects were loaded to four AI programmes commonly used in translating sentences from Turkish to English, DeepL Translate, Google Translate, Magic School and ChatGPT. The translation sentences were analysed with regard to the appropriateness of the translation equivalent by comparing and contrasting the target translation platforms with each other.

The results showed that, even though there exist some correct translations based upon sense-for-sense translation, mostly the programmes fail to capture the intended meaning and provided literal translations rather than the idiomatic ones. In this paper the findings of the study are discussed in details and the pedagogical implications of the study are presented.

Keywords: Idiomatic expressions, machine translation, artificial intelligence, culture, cross-communication

Introduction

The main concern of this study is to discover to what extent online translation programs based on artificial intelligence translate sentences loaded with cultural idiomatic expressions successfully. To find out the answer of this question first of all the background relating machine translation and artificial intelligence is given and then idiomatic expressions are explained with some examples and the findings of the earlier studies are discussed.

²Selma Elyıldırım.

E-mail address: selma.elyildirim@hbv.edu.tr

Literature review

The term ‘artificial intelligence’ (AI) was suggested by John McCarthy, an American computer scientist when a workshop was organized at Dartmouth College in 1956. He used AI as a phrase indicating the act of ‘getting computers to do things humans do easily and without thinking, such as seeing and talking, driving and manipulating objects, as well as planning our everyday lives’ (Wilks, 2019, pp. 7-8). In fact before this, another term ‘Machine Intelligence’ had been coined by Donald Michie to make reference to the use of machine for carrying out certain activities. Wilks (2019) thinks that it was better as the word ‘artificial’ has connotation of trickery and adds that AI develops slowly but steadily. Although there were some hinderances in front of the development of AI, engineers managed to overcome these difficulties and it has increased its expansion in the contemporary world day by day.

AI progress was quite rapid and this progress changed the judgements relating the claims that certain activities such as machine translation (MT) were impossible. A researcher supporting this view was Yehoshua Bar-Hillel and he believed MT system had to possess a large amount of world knowledge to cope with it in 1960. However, with the latest developments in computer science various sources whose size went up to billions have been compiled on the internet and MT has become possible and freely available. Many online dictionaries such as Cambridge, Oxford, Macmillan, Longman, etc. and search engines such as Google, Yandex, Yahoo, etc. provide the service of translation to the computer users.

The integration of machine translation into the process of translating texts from the source language to the target language has brought new changes into the practice. AI has been introduced and incredible applications enabling users to translate texts in seconds have been produced. AI has both increased the speed and the quality of the translation made. Thus individuals have obtained the opportunity of having an efficient cross-cultural communication. Although AI has solved many problems MT faced in transferring meaning from one language to another, still a specific area causes trouble to learners: cultural idioms. As they include cultural features and meanings, their translation requires special care for the conveyance of the message which goes beyond the literal meaning. That is to say, the meaning relating historical, social and cultural contexts of idioms need to be taken into consideration in their translation. AI has a great degree of accuracy for sentences constructed with words referring literal meanings, however it encounters problems when idioms occur in sentences. They do not fit their literal translation and in such cases AI models, which do not possess cultural sensitivity, have difficulty to convey the same message.

Idiomatic Expressions

Idioms which are the expressions whose meaning cannot be deduced from the individual meanings of their components pose great difficulty to machine translation systems. Apart from having their meanings as a unit of words, they have cultural contexts playing a great role in meaning conveyance. For example, idioms such as ‘etekleri zil çalmak’ (being overjoyed) and ‘haber uçurmak’ (send news) make reference to metaphorical or symbolic meanings and hence they are not translated by relying on the word-for-word equivalents. Their literal translation may also lead to misunderstandings due to nonsensical meanings referred.

Empirical studies provide support to the point that idioms are the most difficult items of a language to translate. Although AI systems improve their capacities through machine learning algorithms and access to large bilingual corpora to produce the accurate translation equivalents for cultural idioms, they still struggle with cultural idioms. Brown's study (2023) has shown that AI accuracy goes down to 60 percent when translating sentences containing context and culture specific idioms.

Method

The aim of this study was to figure out to what extent Artificial Intelligence could be used to translate sentences with cultural idioms. In other words, the study focused on whether AI systems translate idiomatic expressions from a source language to the target language without changing the meaning expressed in the source language.

Sample

In order to find out the answers of these questions twenty sentences were constructed with frequently used culturally loaded idiomatic expressions. The idiomatic expressions included in the study was in the form of noun + verb combination and were taken from a Turkish web page presenting the commonly used idiomatic expressions with their meanings (<https://www.kolaykampus.com/blog-detay/en-cok-kullanilan-deyimler-ve-anlamlari>) to have objective criteria for the selection of the items.

5.1. Instrument(s)

After deciding on the expressions the sentences relating the idiomatic meanings they refer to were produced.

5.2. Data collection and analysis

These sentences were written to four AI programmes DeepL Translate, Google Translate, Magic School and ChatGPT. Subsequent to this process, the translated sentences were analysed as regards the equivalence and appropriateness of sentences in the target language. The sentences were checked for equivalence and the translation programmes were evaluated by comparing and contrasting them with each other.

6. Results and Discussion

Twenty sentences constructed from the cultural idiomatic expressions are given with their translations in Appendices. The analyses revealed that except for ChatGPT, three AI platforms translated some of these sentences by making reference to their literal meanings rather than idiomatic meanings they express. In the translation sentences of DeepL Translate, Google Translate this was the case mostly observed. For example, while translating the sentences 'Sevinçten etekleri zil çalıyordu' and 'Müdür öğrencilere aba altından sopa gösterdi', they relied on the literal meaning and gave 'Her skirts were jingling/ringing with joy' and 'The director/manager showed students the stick under the cloak' as translation

equivalents. Magic School had also similar translation equivalents for these sentences, too. However, for Sentence 3 'Durumu idare edinceye kadar akla karayı seçtim', Google Translate provided translation equivalent by considering the sense 'I had a hard time until I managed the situation' whereas Deepl Translate presented the literal translation 'I had to choose between black and white until I managed the situation'. In Sentence 4 'Kadin allem etti kallem etti kocasının aklını çeldi', translation sentences of both Deepl Translate and Google Translate were based on the meaning 'to use all sort of cunning means to reach one's goal'. Magic School was unsuccessful in the translation of the fourth sentence since it produced a translation being away from the intended meaning 'The woman did everything possible to confuse her husband's mind'.

In the translation of Sentence 5 'Dinleyicileri etkilemeyi başardı, itiraf etmeliyim ki ağızı laf yapıyor' all platforms were sucessful at a certain level as they had the expression 'smooth talker' or 'talking well' as the equivalent of 'ağızı laf yapmak'. As to Sentence 6 'Sana mirastan bu kadar düştü, artık bahtına küs' Google Translate gave the expression 'disgrunted' implying 'being disappointed about something' and Deepl Translate had 'being out of luck' as extended senses of the idiom 'bahtına küsmek'. 'Dallanıp budaklandırmak' which means 'expanding and getting complicated' in Turkish was translated with 'no need to get into/drag out/ delve into a long discussion/argument' in Sentence 7. Sentence 8 included the idiom 'tefe koymak' which refers 'to spread rumors' or 'to mock' was translated with 'put someone on the spot/hook' in Deepl and Google but Magic School provided 'turn something against someone' and ChatGPT 'get crucified'. These translation expressions displayed that AI programmes construed the sense of the idiom.

'Göz hapsine almak' in Sentence 9 which means 'watch closely/ keep an eye on someone' was translated correctly in view of the meaning it refers to in all translation programmes. The idiomatic expression under examination in Sentence 10 'Annesi kızına aynı babasının huyunu suyunu aldığı söylüyordu' was 'huyunu suyunu almak' and the phrasal expression 'take after' was preferred by Magic School and ChatGPT; yet Deepl and Google Translate took 'temperament' into account and use 'inherit/absorb father's temperament'. In Sentence 11, the focus was on 'hüsnükuruntuya kapılmak' which implies 'to have a wishful thinking' was translated with the original word 'hüsnükuruntu' by Deepl Translate. Google Translate was confused with the meaning expressed in the sentence and Magic School and ChatGPT had the paraphrase of the idiom as 'having an illusion/ falling into wishful thinking'.

In Sentence 12 'Anlaşmayı imzalamamak için ipe un serip duruyorsun' included 'ipe un sermek' referring the excuses put forward to avoid the thing that should be done was paraphrased as 'putting up obstacles' in Deepl Translate, it was omitted from the sentence in Google Translate. Nevertheless, Magic School had its equivalent 'dragging feet' and ChatGPT 'keep stalling' while translating the sentence. 'kabak tadi vermek' which remarks the negative attitude towards a behaviour was replaced with 'become tiresome' in Deepl Translate and Magic School, 'get on nerves' in Chat GPT and 'take its toll' in Google Translate. 'Nabız yoklamak' having the meaning of understanding the tendency, was converted into 'check someone's pulse' and 'test someone's pulse' with its literal meaning in

DeepL Translate and Google Translate, in turn in Sentence 14. Magic School and ChatGPT considered the sentimental sides with the extended sense and used ‘gauge’ to indicate the evaluation of feelings.

As for Sentence 15, it contained a cultural idiom ‘ocağına incir ağacı dikmek’ which means to ruin one’s family. Apart from ChatGPT, the other translation programmes had real difficulty and gave the literal translation equivalent for the target item. Even Google Translate produced the incorrect translation ‘You’re going to plant my fig tree in our oven?’ for the target sentence due to confusion the idiom caused. The following sentence had ‘pabucunu ters giydirmek’ which refers to put someone into a difficult situation, to swindle. It was translated with its literal meaning by DeepL Translate and Google Translate but Magic School and ChatGPT had paraphrased version of the sentence: ‘A cunning person will turn the tables on you; one must be alert’, and ‘A cunning person can outsmart even the trickiest of folks—you’ve got to stay sharp’, respectively.

The cultural idiom ‘pabucun pahalı olduğunu anlamak’ points out that something is dangerous in Sentence 17. In DeepL, Google Translate and Magic School the idiomatic expression was rendered with literal translation but ChatGPT provided the paraphrase for the idiom as ‘Realizing he was in trouble, he backed off’. In Sentence 18, ‘sıfırı tüketmek’ having the meaning of being broke, going bankrupt was the target idiom, but it was taken from a different perspective and ‘hit rock bottom’ which means to be in an extremely low level was used by ChatGPT and DeepL Translate and the other programmes had the paraphrased version of the translation.

In Sentence 19 ‘üçe beşe bilmek’, not being stingy or not taking the small amounts into account in shopping was the idiom under investigation and it was translated as ‘without worrying/looking at the price/details’. For the last sentence involving the idiom ‘zülfüyara dokunmak’ which means ‘touch someone on sore spot’ was translated in such form in ChatGPT, ‘touch a nerve’ was used by DeepL Translate, but Google Translate and Magic School had the literal meaning in their translation.

Compared with other translation programmes ChatGPT, as can be seen in the sentences given in Appendix E, was more successful than others. It generated the translation equivalences in relation with sense-for-sense approach and presented the idiomatic meanings. This result implied that it searches idioms across various sources and finds the most suitable meaning matching with the target language in translation. Johnson et al. (2022) found that modern AI, if trained properly for the search of extensive datasets, manages to find similar idiomatic expression with 85 percent accuracy. Although Magic School sometimes had the literal translation for the cultural idioms, it mostly provided acceptable translation sentences. When evaluating the translation made by Google Translate and DeepL Translate, they could not cope with cultural idioms as well as the other programmes. Google Translate misunderstood some of the sentences and changed the message given in the original sentence. DeepL Translate based its translation on the literal conversion of meaning rather than the sense conveyed in sentences.

7. Conclusions

The study investigated the effectiveness of AI-based online translation programs in translating sentences with culturally loaded idiomatic expressions. Firstly a brief history of artificial intelligence (AI) and its role in machine translation (MT) were discussed, then the criticism that MT required vast world knowledge was presented. However, advances made in technology, especially in the last decade, and large bilingual corpora created have dispelled the concerns relating MT and eventually it has become widely available through services like Google Translate and DeepL. Additionally it has largely improved speed and overall quality. Although progress has made in MT, cultural idioms remain a persistent challenge as their meanings go beyond literal word meanings. They also refer to historical, social, and cultural contexts. In the light of these points, twenty sentences incorporating frequently used Turkish idioms were translated by using four AI systems: DeepL Translate, Google Translate, Magic School, and ChatGPT. Translations were evaluated for equivalence and appropriateness in the target language. These AI systems were also compared with each other. The results showed that, except for ChatGPT, most platforms often produced literal translations instead of idiomatic equivalents. DeepL and Google mostly rendered idioms with word-for-word English equivalents. Magic School translated several sentences with literal outputs though some translations captured idiomatic senses. ChatGPT successfully provided the translation equivalents for sentences by producing paraphrases that matched idiomatic meanings. This finding implied that ChatGPT had better ability for searching large datasets for culturally equivalent expressions. The conclusion drawn from the study was that modern AI has substantially improved MT but still struggles with culture-specific idioms.

References

Brown, H. (2023). *Early Modern Translation and the Digital Humanities*. J.B. Metzler.

Yorick, W. (2019). Artificial Intelligence: Modern Magic or Dangerous Future? Icon.

Zhang, Y., & Smith, J. (2023). *Title of the translation study or report*.

Dörnyei, Z. (2007a). Creating a motivating classroom environment. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 719-731). New York, NY: Springer.

Dörnyei, Z. (2007b). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Dörnyei, Z. (2012). *Motivation in language learning*. Shanghai: Shanghai Foreign Language Education Press.

Ellis, R. (2012). A lifetime of grammar teaching. *Teaching English in the Two-Year College*, 40(1), 7-19.

Griva, E. (2007). A story-based framework for a primary school classroom. Retrieved July 06, 2009, from: <http://exchanges.state.gov/forum/vols/vol45/no4/p26.htm>.

Kieffer, M. J., & Lesaux, N. K. (2009). Breaking down words to build meaning: Morphology, vocabulary, and reading comprehension in the urban classroom. In M. L. Graves (Ed.), *Essential readings in vocabulary instruction* (pp. 90-101). Newark, DE: International Reading Association.

Ozsoy, G., & Gunindi, Y. (2011). Prospective preschool teachers' metacognitive awareness. *Elementary Education Online, 10*(2), 430-440.

Ravid, R. (2011). *Practical statistics for educators* (4th ed.). Plymouth: Rowman & Littlefield.

Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.

ChatGPT <https://chatgpt.com/>

Magic School AI <https://www.magicschool.ai/>

DeepL Translate <https://www.deepl.com/en/translator>

Appendix A. Turkish sentences translated through AI tools

Sevinçten etekleri zil çalıyordu.

Müdür öğrencilere aba altından sopa gösterdi.

Durumu idare edinceye kadar akla karayı seçtim.

Kadın allem etti kallem etti kocasının aklını çeldi.

Dinleyicileri etkilemeyi başardı, itiraf etmeliyim ki ağızı laf yapıyor.

Sana mirastan bu kadar düştü, artık bahtına küs.

Bu tartışmayı dallanıp budaklandırmaya gerek yok, burada bitirelim.

Olaya ilişkin bir şey söyleme, sonra seni tefe koyarlar.

Altan alta genç kızı göz hapsine almış, ne yaptığıni anlamaya çalışıyordu.

Annesi kızına aynı babasının huyunu suyunu aldığıni söylüyordu.

Genç kız adamın bakışlarından kendine ilgisi olduğu hüsnükuruntusuna kapılıyordu.

Anlaşmayı imzalamamak için ipe un serip duruyorsun.

Kiracısının bu huzursuzlukları artık kabak tadi vermeye başlamıştı.

İşçilerin nabzını yoklayıp maaşa zam oranını belirledi.

Aşçı yamağına bu kadar malzeme konur mu, sen bizim ocağımiza incir ağacımı dikeceksin diye bağıriyordu.

Düzenbaz insana pabucunu ters giydirir, uyanık olmak gerek.

Pabucun pahalı olduğunu anlayınca yanlışından döndü.

Yeniden işe giremeyince sıfırı tüketti.

Üçe beşe bakmadan evin eşyasını satın aldık.

Kahya söyleyecekleri zülfüyara dokunsa da beyden gerçeği gizlemeyecekti.

Appendix B. DeepL Translate

Her skirts were jingling with joy

The director showed students the stick under the cloak.

I had to choose between black and white until I managed the situation.

The woman schemed and plotted to turn her husband's mind.

He managed to impress his audience; I must admit, he's a smooth talker.

That's all you got from the inheritance; now you're out of luck.

There's no need to get into a long discussion about this, let's end it here.

Don't say anything about it, or they'll put you on the spot.

He was secretly watching the young girl, trying to figure out what she was doing.

Her mother said her daughter had inherited her father's temperament.

The young girl was caught up in the hüsnükuruntu that the man's glances meant he was interested in her.

You keep putting up obstacles to avoid signing the agreement.

The tenant's restlessness was starting to become tiresome.

He checked the workers' pulse and determined the wage increase rate.
He yelled, "How can you put so much stuff in the cook's bag? Are you going to plant a fig tree in our kitchen?"

A schemer will make you wear your shoes backwards; you have to be alert.
When he realized the shoes were expensive, he changed his mind.
Unable to get back to work, he hit rock bottom.
We bought the house's belongings without looking at the price.
Even if what the steward said touched a nerve, he would not hide the truth from the master.

Appendix C. Google Translate

Her skirts were ringing with joy.
The manager showed students a stick under her robe.
I had a hard time until I managed the situation.
The woman tried to persuade her husband.
He managed to impress the audience; I must admit, he's a bit of a talker.
You've inherited so much; you're now disgruntled.
There's no need to delve into this argument; let's end it here.
Don't say anything about the incident, or they'll put you on the hook.
He kept a close eye on the young girl, trying to understand what she was doing.
Her mother was telling her daughter that she had absorbed her father's temperament.
The young girl's gaze made her think he was interested in her.
You're just trying to avoid signing the agreement.
Her tenant's unrest was starting to take its toll.
He tested the workers' pulse and determined the rate of the wage increase.
The chef was yelling, "How can you put so much material on your assistant? You're going to plant my fig tree in our oven?"
A cheater can make you wear your shoes backwards; you have to be vigilant.
When he realized the shoe was expensive, he rectified his mistake.
Unable to get back to work, he lost all patience.
We bought the house's furniture without even looking at the details.
Even if what the butler said offended the prince, he wouldn't hide the truth from him.

Appendix D. Magic School

The skirts were ringing with joy
The principal showed students a stick under the robe.
I was in a dilemma until I managed the situation.
The woman did everything possible to confuse her husband's mind.
He managed to impress the listeners; I must admit he can talk well.
You got this much from the inheritance; now curse your fortune.
There's no need to complicate this discussion; let's end it here.
Don't say anything about the incident; they might turn it against you.
He was secretly keeping an eye on the young girl, trying to understand what she was doing.
His mother used to say that her daughter took after her father's character.
The young girl was under the illusion that the man was interested in her based on his looks.
You're dragging your feet to avoid signing the agreement.
The tenant's unrest was starting to become tiresome.
He gauged the workers' feelings and determined the salary increase rate.
The cook was shouting, asking if so much material could be put into the dish, saying, "Are you going to plant a fig tree in our kitchen?"
A cunning person will turn the tables on you; one must be alert.

When he realized the shoe was expensive, he changed his mind.
When he couldn't get back to work, he was completely out of options.
We bought the furniture without looking at the price.
Although the steward touched upon what he would say, he wouldn't hide the truth from the lord.

Appendix E. ChatGPT

She was on cloud nine.
The principal issued a veiled threat to students.
I went through hell trying to manage the situation.
The woman used every trick in the book to persuade her husband.
He managed to impress the audience—I'll admit, he's quite a smooth talker.
This is your share from the inheritance—blame fate if it's not enough.
There's no need to drag out this argument—let's end it here.
Don't say anything about the incident, or you'll get crucified for it.
He was secretly watching the young girl, trying to figure out what she was up to.
Her mother said she had taken after her father in every way.
The young girl was falling into wishful thinking, believing the man's glances meant he liked her.
You keep stalling to avoid signing the agreement.
His tenant's constant issues had really started to get on his nerves.
He gauged the workers' sentiments and then set the raise rate accordingly.
He yelled at the kitchen apprentice, "Do you put this much in? Are you trying to ruin us?!"
A cunning person can outsmart even the trickiest of folks—you've got to stay sharp.
Realizing he was in trouble, he backed off.
Failing to get another job, he hit rock bottom.
We bought the furniture without worrying too much about the price.
Even if his words might touch a sore spot, the steward wouldn't hide the truth from the master.
Authors including an appendix section should do so after References section. Multiple appendices should all have headings in the style used above. They will automatically be ordered A, B, C etc.

AUTHOR BIODATA

Selma Elyıldırım completed her BA degree at the Department of English Language and Literature of Atatürk University in 1990 and became a research assistant at the same department. She did her first MA in the field of TEFL at the Department of Foreign Languages in the Faculty of Education of the same university. She was awarded her second MA and PhD degrees by the Department of Linguistic Studies at Reading University of the United Kingdom. In 2008 she had an assistant professor position at the Discipline of English Language and Literature in the Faculty of Letters at Gazi University. In 2010 she went to the United States of America with an exchange programme between Gazi University and Georgia State University in Atlanta to carry out postgraduate studies. In 2012 she became an associate professor. She has been working as a professor at Ankara Hacı Bayram Veli University. She carried out studies in theoretical and applied linguistics. She edited a book whose title is *Representations of Space in Literature* and wrote a book called *Collocations: Lexical Word Combinations* in 2018. She is acting as an editor in various journals. She is currently the Head of Commission of Forensic Linguistics in the Association of Forensic Studies. Her interest

areas range from semantic-syntax correspondence, language acquisition, psycholinguistics, corpus linguistics, discourse analysis and stylistics to forensic linguistics. She has many articles and book chapters published in international and national journals as well as papers presented in international and national conferences.

School Climate as a Moderator in the Relationship Between Job Satisfaction, Teacher Immunity, and Burnout: A Proposed Moderated Mediation Model

Mehmet TUNAZ ^{a 3}

^a*Nevşehir Hacı Bektaş Veli University, Nevşehir, Türkiye*

APA Citation:

Tunaz, M. (2025). *School Climate as a Moderator in the Relationship Between Job Satisfaction, Teacher Immunity, and Burnout: A Proposed Moderated Mediation Model*. INOED 2025 Conference.

Abstract

Teacher burnout is a growing concern worldwide, with significant implications for teacher well-being, instructional quality, and student outcomes. Understanding the factors that contribute to or mitigate burnout is crucial for enhancing teacher resilience and retention. This study proposes a moderated mediation model that examines how job satisfaction influences burnout among English as a Foreign Language (EFL) teachers. In this model, teacher immunity—a psychological defense mechanism that enables teachers to adapt to professional stressors—is hypothesized to mediate the relationship between job satisfaction and burnout. Moreover, school climate is expected to moderate both the direct and indirect pathways, strengthening or weakening the effects of job satisfaction on teacher immunity and burnout. The study will adopt a cross-sectional survey design with a target sample of 400–600 EFL teachers working in public and private schools across Turkey. Four validated instruments will be employed: the Minnesota Satisfaction Questionnaire to measure job satisfaction, the Teacher Immunity Scale to assess resilience-related mechanisms, the Maslach Burnout Inventory to evaluate burnout, and the Revised School Level Environment Questionnaire to capture perceptions of school climate. Data will be analyzed using Hayes' PROCESS Macro (Models 4 and 8) with bootstrapping procedures to test mediation and moderated mediation effects. It is anticipated that higher job satisfaction will lead to stronger teacher immunity, which in turn will reduce burnout levels. A positive school climate is expected to amplify these effects by fostering collaboration and support. The findings are expected to provide both theoretical and practical insights, guiding educational leaders and policymakers in developing interventions aimed at enhancing teacher well-being, reducing burnout, and improving overall educational outcomes.

Keywords: Job satisfaction, teacher immunity, burnout, school climate, moderated mediation, EFL teachers

Introduction

Teacher burnout has become one of the most pressing issues in educational systems worldwide, with negative implications for teacher well-being, student outcomes, and institutional effectiveness (Maslach et al., 2021). Burnout is characterized by emotional exhaustion, depersonalization, and a reduced sense of personal accomplishment, often resulting in decreased instructional quality and increased teacher turnover (Maslach & Leiter, 2016). Recent research emphasizes the interplay of individual, organizational, and contextual factors in shaping teacher burnout, highlighting the importance of understanding both personal resources and environmental influences (Fathi & Derakhshan, 2019).

Job satisfaction is a critical determinant of teacher retention and performance. Teachers who report higher levels of job satisfaction are more motivated, exhibit higher organizational commitment, and demonstrate greater instructional effectiveness (Skaalvik & Skaalvik, 2017). Conversely, low job satisfaction is strongly correlated with increased stress and burnout (Collie et al., 2012). Teacher immunity, a relatively novel concept in educational psychology, refers to the protective psychological mechanisms teachers develop to cope with occupational challenges (Hiver & Dörnyei, 2017).

³ Corresponding author.

E-mail address: mehmet.tunaz@nevsehir.edu.tr

Understanding how job satisfaction influences teacher immunity and, in turn, affects burnout offers valuable insights into teacher well-being and resilience.

School climate represents the collective perceptions of teachers, administrators, and students regarding the quality and character of school life, including aspects of relationships, practices, and norms (Johnson et al., 2007). A positive school climate has been associated with increased teacher motivation, collaborative practices, and reduced stress levels (Collie et al., 2012). However, the interaction between school climate, job satisfaction, teacher immunity, and burnout remains underexplored, particularly in contexts where English as a Foreign Language (EFL) teachers face unique challenges.

The purpose of this study is to propose a moderated mediation model examining the indirect relationship between job satisfaction and burnout through teacher immunity, with school climate as a moderator. This model aims to answer three main questions: whether teacher immunity mediates the relationship between job satisfaction and burnout, whether school climate moderates the relationship between job satisfaction and teacher immunity, and whether school climate moderates the indirect effect of job satisfaction on burnout through teacher immunity.

Literature review

Job Satisfaction and Teacher Immunity

Job satisfaction reflects teachers' overall affective orientation toward their work (Weiss et al., 1967). High job satisfaction is associated with better emotional regulation, enhanced self-efficacy, and adaptive coping mechanisms (Klassen & Chiu, 2010). Teacher immunity, as defined by Hiver (2017), is a psychological defense system that enables teachers to manage and adapt to professional stressors. Research suggests that supportive organizational conditions and positive affective experiences, such as job satisfaction, can strengthen teacher immunity by fostering resilience and growth-oriented mindsets (Hiver & Dörnyei, 2017). In this context, job satisfaction is considered a critical factor in enhancing teachers' internal resources to cope with challenges. Therefore, it is expected that job satisfaction will positively predict teacher immunity.

Teacher Immunity and Burnout

Teacher immunity plays a protective role against burnout by buffering the impact of stress and promoting adaptive coping strategies (Derakhshan et al., 2023). Teachers with strong immunity are more likely to maintain motivation, persistence, and emotional well-being despite challenges such as student misbehavior or administrative pressures (Fathi & Derakhshan, 2019). Conversely, teachers with low immunity tend to be more vulnerable to stressors and may exhibit maladaptive coping mechanisms, ultimately leading to increased levels of burnout. This suggests that teacher immunity functions as a crucial psychological resource in preventing or reducing burnout among teachers. Accordingly, it is hypothesized that higher levels of teacher immunity will be associated with lower levels of burnout.

Job Satisfaction and Burnout

Previous research consistently demonstrates a negative relationship between job satisfaction and burnout (Skaalvik & Skaalvik, 2017; Maslach & Leiter, 2016). Teachers who are satisfied with their work environment, relationships with colleagues, and professional growth opportunities are more resilient to stressors and less prone to experiencing emotional exhaustion. In contrast, dissatisfaction in these areas tends to exacerbate feelings of depersonalization and reduced personal accomplishment.

Thus, job satisfaction acts as a protective factor, reducing the risk of burnout. Based on this evidence, it is anticipated that job satisfaction will negatively predict burnout among teachers.

Mediating Role of Teacher Immunity

The relationship between job satisfaction and burnout may not be direct but rather operate through intermediate psychological mechanisms. According to the Conservation of Resources Theory (Hobfoll, 2018), individuals strive to acquire, retain, and protect valued resources. Job satisfaction can be conceptualized as a personal resource that enhances teachers' capacity to develop and maintain teacher immunity. This, in turn, equips them with the resilience necessary to cope with occupational stressors and prevent burnout. Therefore, teacher immunity is expected to serve as a mediator in the relationship between job satisfaction and burnout, explaining how job satisfaction influences burnout outcomes.

Moderating Role of School Climate

School climate encompasses the collective social, academic, and organizational structures within a school setting (Johnson et al., 2007). A positive climate characterized by collaboration, support, and innovation can amplify the beneficial effects of job satisfaction on teacher immunity by creating an environment where teachers feel valued and empowered. Conversely, a negative climate marked by conflict or lack of support may weaken this relationship, hindering teachers' ability to build strong immunity even when they are generally satisfied with their job (Polat & İskender, 2018). Furthermore, school climate may also influence the strength of the indirect effect of job satisfaction on burnout through teacher immunity, suggesting the presence of a moderated mediation model (Hayes, 2017). This implies that the degree to which teacher immunity mediates the job satisfaction–burnout relationship depends on the level of school climate. In schools with highly positive climates, the protective role of teacher immunity is expected to be stronger, whereas in schools with negative climates, this mediating effect may be diminished.

Proposed Research Model

The proposed conceptual model positions job satisfaction as an antecedent variable influencing teacher immunity, which subsequently affects burnout. School climate is hypothesized to moderate the relationship between job satisfaction and teacher immunity, as well as the indirect pathway linking job satisfaction to burnout through teacher immunity. This model integrates both mediation and moderation processes, providing a comprehensive framework for understanding the complex interplay among these variables.

Method

Participants

The target population of this study consists of English as a Foreign Language (EFL) teachers working in public and private schools across Turkey. Considering recommendations for structural equation modeling, a minimum sample size of 400 participants is proposed to ensure sufficient statistical power (Kline, 2015). Stratified sampling will be employed to capture variability across different school types, regions, and demographic characteristics.

7.1. Instrument(s)

Four validated scales will be utilized to measure the study variables. Job satisfaction will be assessed using the Minnesota Satisfaction Questionnaire (MSQ) developed by Weiss et al. (1967), which includes 10 items rated on a five-point Likert scale. Teacher immunity will be measured using the Teacher Immunity Scale created by Hiver (2017), consisting of 16 items rated on a six-point Likert scale. Burnout will be evaluated using the Maslach Burnout Inventory (Maslach et al., 1996), comprising five items rated on a seven-point Likert scale. Finally, school climate will be assessed using the Revised School Level Environment Questionnaire developed by Johnson et al. (2007), which includes eight items rated on a five-point Likert scale.

7.2. Data collection procedures

Ethical approval will be sought prior to the commencement of data collection. Participation will be entirely voluntary, and informed consent will be obtained from all participants. To ensure broad participation, surveys will be distributed both online and in paper format. Anonymity and confidentiality will be maintained throughout the study, with all data used solely for research purposes.

7.3. Data analysis

Data analysis will proceed in several stages. Preliminary analyses will include descriptive statistics, normality checks, and reliability assessments using Cronbach's alpha to ensure internal consistency of the scales. Mediation will be tested using PROCESS Macro Model 4 (Hayes, 2017), while moderated mediation will be examined using Model 8. Bootstrapping with 5000 resamples will be employed to estimate the significance of indirect effects, and statistical significance will be set at the conventional level of $p < .05$.

8. Results

It is anticipated that the results will provide strong support for the hypothesized model. Job satisfaction is expected to positively predict teacher immunity, indicating that teachers who are more satisfied with their work environment, professional relationships, and job characteristics will exhibit higher levels of psychological resilience and adaptive coping mechanisms. Teacher immunity, in turn, is expected to negatively predict burnout, suggesting that teachers with stronger immunity will experience lower levels of emotional exhaustion and depersonalization. Moreover, job satisfaction is likely to demonstrate a negative direct relationship with burnout, confirming its role as a protective factor against stress-related outcomes.

The mediating role of teacher immunity is expected to be significant, demonstrating that the effect of job satisfaction on burnout operates partly through the enhancement of teachers' psychological resources. In other words, when teachers are satisfied with their jobs, they are more likely to build strong immunity, which subsequently reduces their likelihood of experiencing burnout.

School climate is expected to moderate both the direct and indirect pathways within the model. In schools characterized by a positive climate, the relationship between job satisfaction and teacher immunity is anticipated to be stronger, reflecting the supportive and collaborative nature of the environment. Similarly, the indirect effect of job satisfaction on burnout via teacher immunity is expected to be amplified in such contexts, illustrating the crucial role of environmental factors in shaping teachers' psychological well-being. Conversely, in schools with negative or unsupportive

climates, these relationships are expected to be weaker, highlighting the potential risks associated with adverse organizational conditions.

Number tables consecutively in accordance with their appearance in the text. Place footnotes to tables below the table body and indicate them with superscript lowercase letters. Avoid vertical rules. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article. Figures should have the same style as tables; however, the title of a figure should appear below it (see Figure 1).

9. Discussion

The proposed study seeks to integrate personal, organizational, and contextual variables into a comprehensive framework for understanding teacher burnout. By demonstrating the mediating role of teacher immunity and the moderating role of school climate, this research can contribute to the development of holistic intervention strategies. If the hypotheses are supported, the findings will suggest that efforts to enhance teacher well-being should address not only individual-level factors, such as job satisfaction, but also the broader organizational environment.

This model emphasizes the interconnectedness of teachers' internal psychological resources and external organizational support systems. For example, even when teachers are highly satisfied with their work, they may still be at risk of burnout if they lack sufficient immunity or work in a toxic climate. Conversely, a positive school climate can amplify the benefits of job satisfaction by fostering collaboration, innovation, and shared purpose.

10. Limitations and future directions

The proposed study has several limitations that should be considered. As a cross-sectional design, it limits the ability to draw causal inferences. Future research should employ longitudinal or experimental designs to establish temporal relationships among the variables. Additionally, the cultural context of the study may affect the generalizability of the findings, as the model is tested within the specific context of Turkish EFL teachers. Replicating the study in different countries or educational settings would help determine the robustness of the results. Another limitation involves the reliance on self-report data, which may be subject to biases such as social desirability or common method variance. Future studies could include additional data sources, such as classroom observations or administrator evaluations, to triangulate findings.

References

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social-emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology, 104*(4), 1189–1204. <https://doi.org/10.1037/a0029356>

Derakhshan, A., Wang, Y., Wang, Y., & Ortega Martín, J. L. (2023). Towards innovative research approaches to investigating the role of emotional variables in promoting language teachers' and learners' mental health.

Fathi, J., & Derakhshan, A. (2019). Teacher self-efficacy and emotional regulation as predictors of burnout among English language teachers. *Iranian Journal of Language Teaching Research*, 7(2), 37–54.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hiver, P. (2017). Tracing the signature dynamics of language teacher immunity: A retrodictive qualitative modeling study. *The Modern Language Journal*, 101(4), 669–690.

Hiver, P., & Dörnyei, Z. (2017). Language teacher immunity: A double-edged sword. *Applied Linguistics*, 38(3), 405–423.

Johnson, B., Stevens, J. J., & Zvoch, K. (2007). Teachers' perceptions of school climate: A validity study of scores from the Revised School Level Environment Questionnaire. *Educational and Psychological Measurement*, 67(5), 833–844.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach Burnout Inventory manual* (3rd ed.). Consulting Psychologists Press.

Maslach, C., & Leiter, M. P. (2016). Understanding the burnout experience: Recent research and its implications for psychiatry. *World Psychiatry*, 15(2), 103–111.

Polat, D. D., & İskender, M. (2018). Exploring teachers' resilience in relation to job satisfaction, burnout, organizational commitment and perception of organizational climate. *International Journal of Psychology and Educational Studies*, 5(1), 1–13.

Skaalvik, E. M., & Skaalvik, S. (2017). Motivated for teaching? Associations with school goal structure, teacher self-efficacy, job satisfaction and emotional exhaustion. *Teaching and Teacher Education*, 67, 152–160.

Weiss, D. J., Dawis, R. V., England, G. W., & Lofquist, L. H. (1967). *Manual for the Minnesota Satisfaction Questionnaire*. Minnesota Studies in Vocational Rehabilitation.

AUTHOR BIODATA

Mehmet Tunaz is an Assistant Professor in the English Translation and Interpretation Department at Nevşehir Hacı Bektaş Veli University, a position he has held since 2021. Previously, he served as an English language teaching instructor at Erciyes University, School of Foreign Languages. He completed his PhD at Erciyes University, English Language and Literature Department. He obtained his MA degree in 2014 and his BA degree in 2010, both from Çukurova University, English Language Teaching Department. His research interests include program evaluation and development for language teaching, language teaching methodologies, teacher education and development, curriculum design, and assessment/evaluation in language teaching."

Conjunctive Devices in Undergraduate Literature Reviews A Corpus-Based Study

Fatma Çiloglan Konur ^a ⁴

^a Assistant Professor, Gaziantep, Türkiye

APA Citation:

Çiloglan Konur, F. (2025). Conjunctive devices in undergraduate literature reviews a corpus-based study.

Abstract

This corpus-based study investigates the use of conjunctive devices in literature reviews written by second-year undergraduate learners in an English Language Teaching program. The study examines how 57 EFL students employ additive, adversative, causal, and temporal connectors, drawing on Halliday and Hasan's (1976) taxonomy of cohesive conjunctions, to construct cohesion in their academic writing. A 72,000-word corpus was compiled from students' graded literature review assignments and analyzed using #LancsBox for frequency, normalized distribution, and contextual patterns. Findings reveal a strong overreliance on additive devices—especially *and*—which accounted for more than half of all instances. Causal and adversative connectors appeared less frequently, with students favoring basic forms such as *because*, *so*, and *but* over more formal alternatives like *therefore* or *however*. Temporal devices were used minimally. This reflects the thematic rather than sequential organization of literature reviews. Qualitative concordance analysis further indicated occasional misuse, redundancy, and sentence-initial placement of coordinating conjunctions. Overall, the results suggest that while learners can produce basic cohesion, their conjunctive repertoire remains limited, signaling a need for targeted instruction to develop rhetorical awareness and enhance academic writing proficiency.

Keywords: conjunctive devices; cohesion; corpus-based analysis; EFL academic writing; literature review writing

Introduction

Cohesion plays a central role in academic writing. It enables readers to follow arguments, identify relationships between ideas, and interpret the overall flow of discourse. Among the various cohesive elements, conjunctive devices are particularly important. They signal logical, contrastive, causal, and temporal relations within and across sentences. For undergraduate learners of English as a foreign language (EFL), mastering these connectors is essential for producing coherent and academically appropriate texts, however, research consistently indicates that learners often struggle with their accurate and varied use. The present study examines the use of conjunctive devices in literature reviews written by undergraduate students in an English Language Teaching (ELT) program. The aim is to document how learners employ additive, adversative, causal, and temporal connectors when constructing academic arguments, and also to identify patterns that may reflect developmental challenges in their academic writing. By analyzing a corpus of student-produced texts, this study seeks to contribute to a clearer understanding of learners' cohesive practices and the pedagogical implications for writing instruction.

Literature review

Cohesion is a central feature of effective academic writing. Since it enables readers to follow arguments logically and understand how ideas relate to one another. Halliday and Hasan (1976) define cohesion as the semantic ties that hold a text together, encompassing reference, substitution, ellipsis, lexical cohesion, and conjunctions. Among these categories, conjunctive devices function as key

⁴ Corresponding author.

E-mail address: author@affiliation.xxx.xx

semantic markers that signal additive, adversative, causal, and temporal relations. Thereby contributing directly to coherence. These logical and rhetorical functions form the foundation of the present study's analytical focus.

In academic discourse, conjunctive devices play an essential role in linking propositions, clarifying arguments, and enhancing textual readability. As Sanosi (2024) notes, conjunctions operate at both grammatical and semantic levels, which allows writers to express relationships such as similarity, contrast, cause-effect, and sequence. However, EFL learners frequently misuse or underuse these devices, resulting in less coherent and sometimes fragmented texts. Such difficulties may diminish the overall rigor of students' academic writing and, as a result, hinder reader comprehension.

A substantial body of research highlights recurrent challenges in EFL learners' use of conjunctive devices. Studies show a tendency to overuse high-frequency additive connectors, *and*, while underusing more advanced or formal alternatives such as *however* and *therefore* (Granger & Tyson, 1996; Sayah & Fatemi, 2013; Özçelik, 2023). Sanosi's (2024) comparative corpus-based analysis further demonstrates that learners often rely heavily on simple additive forms. They insufficiently employ adversative and causal connectors. These patterns suggest limited lexical range and a lack of awareness regarding the functional diversity of conjunctions.

Additionally, learners frequently employ conjunctions in syntactically inappropriate ways, such as beginning sentences with *And* or *But*. Although this practice is increasingly accepted in informal registers, it remains discouraged in formal academic writing (Chang & Swales, 1999; Tocalo et al., 2022). Nevertheless, Sanosi (2024) notes that even native academic writers occasionally position conjunctions at the start of sentences, indicating a gradual shift toward more flexible stylistic norms.

Halliday and Hasan's (1976) taxonomy offers a comprehensive framework for examining conjunctive relations. The categorization is as follows; additive, adversative, causal, and temporal. Due to its clarity and applicability to both quantitative and qualitative analyses, this model remains widely used in discourse and corpus studies (Hyland & Jiang, 2017; Hosseinpur & Pour, 2022). It provides a systematic basis for assessing not only frequency but also the functional deployment of cohesive devices across texts.

Corpus-based research has been instrumental in providing empirical insights into learners' use of conjunctions. Meyer (2023) highlights the strength of corpus methods in producing reliable, data-driven descriptions of language patterns. Sanosi's (2024) study, for example, compares a 125,000-word Arab EFL learner corpus (ALAWEC) with the BAWE reference corpus, using tools such as AntConc and #LancsBox to measure frequency and distribution of conjunctive devices. Similar corpus-based investigations (e.g., Liu, 2021; Kashiha, 2022) consistently report limited variety and overreliance on high-frequency connectors in learner writing, emphasizing the need for pedagogical support aimed at improving learners' functional and contextually appropriate use of cohesion.

Research questions

To guide the investigation into how undergraduate EFL learners employ conjunctive devices in their academic writing, the study addresses the following research questions:

1. How frequently and in what ways do undergraduate EFL students use the four categories of conjunctive devices—additive, adversative, causal, and temporal—in their literature reviews?
2. What patterns of overuse, underuse, or misuse can be identified in students' deployment of conjunctive devices, and what do these patterns reveal about their developing academic writing proficiency?

Method

Sample / Participants

The sample consisted of 57 second-year undergraduate students enrolled in the English Language Teaching (ELT) department of a private university in Turkey during the Fall semester of the 2024–2025 academic year. All participants were non-native speakers of English and had previously received instruction in academic writing through required coursework in their program. The students were selected through convenience sampling, as they constituted an intact class completing the same course requirements. Their ages typically ranged from 19 to 21, though specific demographic characteristics such as gender or GPA were not central to the purpose of this study. The writing samples used for analysis were collected from each student’s final graded assignment in an Academic Writing course. Prior to data collection, all papers were anonymized to ensure confidentiality and protect students’ identities.

10.1. Instrument(s)

The primary instrument for this study was the corpus compiled from students’ written literature review assignments. These literature reviews, ranging from 1,000 to 2,500 words, served as authentic academic writing samples reflecting learners’ natural use of conjunctive devices. In addition, the study employed #LancsBox (Version X), a corpus analysis software developed by Lancaster University, to conduct frequency counts, concordance searches, and keyword-in-context (KWIC) analysis. #LancsBox has been widely used in corpus-based linguistic research and provides reliable tools for identifying lexical patterns, though reliability and validity are inherent to the software’s established algorithms rather than researcher-developed measures. No additional questionnaires, scoring instruments, or tests were used.

10.2. Data collection procedures

Data collection took place at the end of the semester when students submitted their final literature review assignment titled “The Role of Motivation in Second Language Acquisition.” Students were required to write a 4–10 page literature review incorporating at least 15 academic sources published within the last decade and to follow APA style guidelines. These assignments were originally produced for course assessment purposes, and only after submission were they compiled for research analysis with the students’ informed consent.

All texts were exported in plain-text format and organized into a unified corpus totaling approximately 72,000 words. The researcher prepared the corpus for analysis by removing identifying information and ensuring consistent formatting. Ethical approval for data collection and use was obtained from the university’s Institutional Ethics Committee, and participation was voluntary. The timetable for the study included corpus compilation (Week 1), preliminary cleaning (Week 2), and analysis (Weeks 3–6).

10.3. Data analysis

Data analysis followed a corpus-based descriptive approach. Using #LancsBox, the researcher conducted frequency analyses to identify all instances of explicit conjunctive devices across the corpus. These results were then normalized per 1,000 words to account for variations in text length. The KWIC concordance tool was used to examine the contextual use, syntactic positioning, and functional role of each conjunction. Conjunctive devices were categorized according to Halliday and

Hasan's (1976) taxonomy, which includes four major types: additive, adversative, causal, and temporal.

Following initial categorization, patterns such as overuse, underuse, redundancy, or misuse were documented qualitatively. To ensure reliability, a second rater independently coded 20% of the dataset. Inter-rater reliability was calculated using Cohen's Kappa, resulting in a coefficient of 0.87, indicating strong agreement. The analysis thus combined quantitative frequency measures with qualitative interpretation to provide a comprehensive understanding of students' conjunctive device use.

11. Results

This section presents the findings of the corpus-based descriptive analysis of conjunctive devices in second-year undergraduate ELT students' literature reviews. The analysis focused on four categories of conjunctions based on Halliday and Hasan's (1976) framework: additive, adversative, causal, and temporal. Frequency counts, normalized per 1,000 words, and contextual patterns are reported, followed by illustrative concordance lines extracted using #LancsBox.

Overall Use of Conjunctive Devices

The analysis revealed a total of **1,203** instances of explicit conjunctive devices across the student corpus. When normalized per 1,000 words, the overall frequency of conjunctive devices was **16.71**. As shown in Table 1, additive conjunctions were the most frequently used category, followed by causal and adversative devices. Temporal conjunctions were the least frequent.

Table 1: Frequency and Distribution of Conjunctive Devices

Conjunctive Category	Raw Frequency	Normalized (per 1,000 words)	Percentage of Total
Additive	678	9.42	56.4%
Causal	229	3.18	19.0%
Adversative	202	2.81	16.8%
Temporal	94	1.30	7.8%
Total	1,203	16.71	100%

These findings indicate a clear preference for additive relations, with more than half of all conjunctions falling into this category. The distribution is consistent with previous studies that document learner overreliance on additive devices (Sanosi, 2024; Sayah & Fatemi, 2013).

Most Frequent Conjunctive Devices

Table 2 displays the ten most frequently used individual conjunctive devices. "And" was the most commonly used, accounting for over one-third of all items. The next most frequent were "because," "but," "also," and "so." Other formal connectors such as "moreover," "however," and "therefore" occurred less frequently.

Table 2: Top 10 Conjunctive Devices by Frequency

Rank	Conjunctive Device	Raw Frequency	Normalized Frequency
1	and	437	6.07
2	because	121	1.68
3	but	97	1.35
4	also	81	1.12
5	so	72	1.00
6	however	55	0.76
7	in addition	42	0.58
8	then	36	0.50
9	therefore	31	0.43
10	finally	27	0.37

The heavy reliance on “and” suggests a limited variety in student writing and a tendency to use basic connectors repetitively. While formal adversative and causal conjunctions were present, their frequency was relatively low, reflecting restricted lexical range.

Distribution by Conjunctive Type

Additive Devices

Additive devices comprised 56.4% of all conjunctions. “And” was the most dominant, often used in sentence-initial position or to join long clauses. Other items such as “also” and “in addition” were used less frequently and occasionally redundantly (e.g., “and also”).

Examples from concordance:

- “**And** this increases their motivation in class.”
- “It is important, **also**, for teacher feedback.”
- “**In addition**, learners enjoy the classroom activities.”

Causal Devices

Causal conjunctions accounted for 19% of the data. The most common were “because” and “so,” typically used to explain learner behaviors or motivation. The connector “therefore” was used more sparingly and sometimes inappropriately, signaling uncertainty in function.

Examples:

- “They are motivated **because** of rewards.”
- “The teacher gives feedback, **so** students feel more confident.”
- “It is important, **therefore**, to encourage learners.”

Adversative Devices

Adversative devices represented 16.8% of usage. “But” and “however” were the most frequent, while others like “on the other hand” were rare. Sentence-initial “but” was common, despite traditional academic writing conventions discouraging its use in this position.

Examples:

- “**But** not all learners feel the same way.”
- “**However**, this may not work for every student.”
- “**On the other hand**, extrinsic motivation can fail.”

Temporal Devices

Temporal connectors were the least frequent, comprising only 7.8% of the total. “Then” and “finally” appeared primarily in the organizational structure of the literature reviews. Their limited usage suggests that students structured their texts thematically rather than sequentially.

Examples:

- “**First**, theories are explained.”
- “**Then**, classroom strategies are discussed.”
- “**Finally**, cultural differences are explored.”

12. Discussion

The results of this study provide meaningful insight into how second-year undergraduate EFL learners construct cohesion in academic literature reviews through conjunctive devices. Rather than repeating the findings, this section interprets their significance in light of the study’s aims. The dominance of additive connectors—especially *and*—indicates that students rely heavily on familiar cohesive resources when navigating cognitively demanding writing tasks. This pattern is consistent with previous observations that learners often depend on high-frequency additive forms (Sanosi, 2024; Sayah & Fatemi, 2013; Granger & Tyson, 1996), yet the present study extends this understanding by showing how such reliance manifests within the specific genre of literature reviews.

The limited use of adversative and causal devices suggests that students may not yet possess the rhetorical flexibility needed to articulate contrast, evaluation, and reasoning—key components of literature review writing. Although underuse of these connectors has been documented among EFL learners elsewhere (Özçelik, 2023; Kashiha, 2022), the current findings highlight how this limitation may influence the structure and critical depth of students’ academic discourse. It is possible that students’ choices reflect a developing awareness of genre expectations rather than a complete lack of knowledge. Given that the participants were only in their second year of study, their restricted repertoire may simply reflect their stage of academic literacy development (Hyland & Jiang, 2017; Meyer, 2023).

Several factors may explain why more advanced conjunctions—such as *however*, *therefore*, or *consequently*—appeared less frequently or were sometimes used inappropriately. One explanation is that the cognitive load of synthesizing multiple sources may have led students to prioritize idea generation over more sophisticated cohesive structuring. Another plausible explanation is that students may have transferred patterns from informal or spoken registers, where sentence-initial *and* or *but* is increasingly accepted (Chang & Swales, 1999; Tocalo et al., 2022). Although Sanosi (2024) notes that even native academic writers sometimes use sentence-initial conjunctions, EFL students may be uncertain about when such usage is stylistically appropriate.

While the findings align largely with expectations based on existing literature, they also raise questions about instructional context. It is possible that the course’s emphasis on content and source integration overshadowed explicit instruction on cohesive devices. Alternatively, students may not have fully internalized feedback from prior writing tasks, or the time available for the assignment may

have been insufficient for deeper revision. These potential explanations point to instructional and contextual factors rather than learner deficits alone.

13. Conclusions

This study contributes to a clearer understanding of how undergraduate EFL learners use conjunctive devices when composing academic literature reviews. The findings show that while students demonstrate an ability to produce basic cohesion, their reliance on high-frequency additive connectors and their limited use of adversative, causal, and temporal devices suggest that their cohesive control remains at an early developmental stage. These patterns highlight the need for more explicit instruction on the rhetorical functions of conjunctions and opportunities for students to practice applying a wider range of cohesive strategies in academic writing.

The study concludes that supporting learners in developing a more varied and contextually appropriate repertoire of conjunctive devices may enhance the coherence, critical depth, and academic tone of their writing. Future research may investigate how targeted instruction or scaffolded practice influences learners' conjunctive choices over time.

References

Chang, Y., & Swales, J. (1999). Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? *TEXT - Interdisciplinary Journal for the Study of Discourse*, 19(2), 253–277.

Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17–27.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hosseinpur, R., & Pour, Z. M. (2022). The use of cohesive devices in EFL learners' argumentative writing: A corpus-based approach. *Language Teaching Research Quarterly*, 25, 12–32.

Hyland, K., & Jiang, F. K. (2017). Is academic writing becoming more informal? A diachronic study of academic phraseology. *English for Specific Purposes*, 45, 40–51.

Kashiha, H. (2022). Investigating cohesive devices in EFL learners' academic writing: A comparative study. *Journal of English for Academic Purposes*, 55, 101072.

Liu, Q. (2021). Cohesion in Chinese EFL learners' academic writing: A corpus-driven study. *International Journal of Applied Linguistics*, 31(3), 354–370.

Meyer, C. F. (2023). *Corpus linguistics and linguistic theory: Advances and applications*. Cambridge University Press.

Özçelik, A. (2023). Cohesive devices in Turkish EFL learners' writing: A corpus-based analysis. *TESOL Journal*, 14(1), e00288.

Sanosi, A. B. (2024). Revisiting cohesion in academic writing: A corpus-based study of conjunctive devices in EFL learners' texts. *Asian Journal of English Language Teaching*, 34(1), 25–48.

Sayah, L., & Fatemi, M. A. (2013). A study of the use of conjunctive adjuncts in Iranian EFL learners' writing. *Theory and Practice in Language Studies*, 3(7), 1216–1222.

Tocalo, M. G., et al. (2022). The positioning of discourse markers in EFL academic writing: A diachronic study. *Linguistics and Education*, 69, 101037.

AUTHOR BIODATA

Fatma Çiloğlan Konur is an Assistant Professor in the English Language Teaching Department at Hasan Kalyoncu University. She completed her BA in English Language and Literature at Gaziantep University, followed by an MA in English Language Education at Gaziantep University and a PhD in English Language Education at Çukurova University.

Her academic work mainly focuses on English language education, academic writing, intercultural sensitivity, teacher education, and corpus linguistics. She continues to conduct research and teach in these areas.

EFL learners' frequency and perceptions of using ChatGPT in informal digital learning of English: Does personality matter?

Zeynep Daşer ^{a 5}, Bahar Kar ^b

^a School of foreign languages, Ataturk University, Erzurum, Türkiye

^b School of foreign languages, Ataturk University, Erzurum, Türkiye

APA Citation:

Daşer, Z., & Kar, B. (2025). EFL learners' frequency and perceptions of using ChatGPT in informal digital learning of English: Does personality matter?

Abstract

The surge in popularity of cutting-edge artificial intelligence tools, namely ChatGPT, has been noteworthy in recent years. However, the perceptions of students regarding the use of ChatGPT for informal digital learning of English (IDLE) has been neglected amid debates concerning its potential and challenges. This study employed quantitative methodology (cross-sectional survey design) to scrutinize students' perceptions of using ChatGPT for IDLE along with their personality profiles and to unearth the possible relationship between the two variables. The subjects were 535 English majors at a state university in Türkiye. Data were gathered by using the Big Five Inventory and the Technology Acceptance Model (TAM) questionnaire and analyzed using descriptive and inferential statistics. Results suggested that a majority of participants used ChatGPT for IDLE, but not very often. They had positive perceptions about ChatGPT regarding all five dimensions of the TAM. The most prevailing personality trait was openness to experience, while neuroticism was the least dominant one among participants. A statistically significant association was found between ChatGPT usage frequency and extraversion, agreeableness, and openness traits. Except for neuroticism, all personality traits exhibited significant relationships with students' perceptions, highlighting the need for customizing ChatGPT integration in the teaching practicum.

Keywords: ChatGPT; informal digital learning of English; personality traits, students' perceptions; the technology acceptance model

Introduction

The time frame following 2010 in the realm of computer-assisted language learning (CALL) has been deemed as *ecological CALL* due to the rise in popularity of mobile digital devices, which have expanded the learning ecology by fostering autonomous second language (L2) acquisition beyond the walls of traditional classrooms (Chun, 2019). Within the framework of ecological CALL, the phenomenon known as *Informal Digital Learning of English* (IDLE), which refers to "self-directed English activities in informal digital settings, motivated by personal interests and undertaken independently without being assessed by a teacher" (Lee & Lee, 2021, p. 359), has gained recognition among digital natives worldwide (Lee & Drajati, 2019). In this regard, a growing body of research has demonstrated that encouraging students' IDLE can improve their language proficiency and yield positive psychological effects (Lee & Drajati, 2019). Notably, IDLE intersects with large language models (LLMs) such as Chat Generative Pre-trained Transformer (ChatGPT)-3.5 and ChatGPT-4.0 by OpenAI, which can anticipate, understand, and build sentences, generate translations, provide explanations of vocabulary and grammar, and suggest style improvements in a human-like manner, thereby promoting L2 learning.

⁵ Corresponding author.

E-mail address: author@affiliation.xxx.xx

Several empirical studies attempted to investigate the potential benefits and challenges of using ChatGPT in general education (Rudolph et al., 2023), and language education, in particular (Kovačević, 2023). Gaining in-depth knowledge of this technology is essential for teachers to make informed decisions about leveraging this chatbot for educational purposes. Nevertheless, there is a dearth of scholarly research examining how L2 learners accept and use ChatGPT (e.g., Liu & Ma, 2023) for learning English outside the classroom. To fill that void in the literature, the present study is underpinned by the *technology acceptance model* (TAM) created by Davis (1989) and is based on the notion of IDLE in the extramural digitalized learning ecology (Lee & Drajati, 2019) to scrutinize perceptions of students who learn English as a foreign language (EFL) and their use of ChatGPT for learning English beyond the classroom.

Empirical research vis-à-vis learners' IDLE and their individual differences (IDs) is scarce (Lee & Drajati, 2019; Lee et al., 2021), whereas IDLE research has almost become widespread. In a similar vein, studies on the interplay between personality and technology acceptance (Barnett et al., 2015) or ChatGPT use (De Winter et al., 2023; Li, 2023) are limited and have contrasting outcomes. Informed by the *Big Five Personality Traits Model* (McCrae & Costa, 2003) and driven by this gap, this research seeks to examine the possible correlation between EFL learners' personality profiles and their acceptance and use of ChatGPT in IDLE. The findings of this study may enhance our understanding of IDLE and personality, as well as antecedents of ChatGPT use, contributing to close the disciplinary gap that exists between psychology, L2 acquisition, and CALL.

Literature review

ChatGPT and Informal Digital Learning of English

“Perhaps, after all, we ‘learn’ language in the same way that an animal ‘learns’ the forest, or a plant ‘learns’ the soil” (Van Lier, 2000, p. 259). Van Lier’s (2000) assertion reminds the importance of language learning in natural, out-of-class settings, which is facilitated by ease of access to modern technology. Against this backdrop, there has been a mushrooming of studies on language learning beyond the classroom (LBC) (Reinders & Benson, 2017). This line of inquiry has emerged in two subareas: LBC offline and online. Online LBC, particularly in the form of IDLE, is on the rise in EFL settings where students do not have much chance to be exposed to English in their daily lives, as Soyoof et al. (2021) revealed. Based on Benson’s (2011) seminal framework, IDLE has four dimensions: “location, formality, pedagogy, locus of control” (Benson, 2011, p. 8). Research on IDLE (Lee et al., 2021) demonstrated that EFL learners both receive and produce content in a variety of informal digital settings, which promote autonomy, increase motivation, create minimal or no anxiety, and offer several psychological and emotional gains.

Reaching 100 million active users just after two months it was introduced in 2018, ChatGPT has become the most rapidly expanding application ever (ChatGPT sets record for fastest growing user base - analyst note, 2023). This popular cutting-edge LLM has numerous advantages for language learners as it engages them in real-time conversations, provides prompt feedback for their writing, offers individualized learning opportunities, and more (Kostka & Toncelli, 2023). Despite the extensive use of this state-of-the-art technology, little is known about how particular user groups such as EFL learners engage with and profit from it (Ma & Huo, 2023); therefore, further investigation is necessary. Scrutinizing learners’ perceptions and opinions regarding the use of ChatGPT for self-directed L2 learning is of utmost importance for several reasons. Initially, unveiling information from learners’ first-hand encounters might provide insights about the extent to which ChatGPT meets student needs and expectations to maximize the efficacy of this popular LLM in instructional

practices. Determining learner perceptions can help program developers to make certain adjustments in the ChatGPT interface to make it more user-friendly based on user experience. Feedback from users may inform policy makers and language teachers in integrating the tool into the curriculum, as well. By investigating learner perceptions towards ChatGPT, the present research aims to obtain data that might inform the effective application of this technology that will meet individual learner needs. Indeed, there are a few empirical studies that aim to hear students' voices on utilizing ChatGPT for their L2 learning, but they approach the construct in a purely technical manner, failing to address the psychological factors. One timely study that examines how students perceive and use ChatGPT in naturalistic settings is worth mentioning. Underpinned by the TAM and IDLE frameworks, Liu and Ma's (2023) research employed a cross-sectional survey design and sampled 450 mostly Chinese EFL undergraduate and postgraduate learners as well as those who were working or in their gap year. They developed a questionnaire, which was also used in the present study, by adopting various TAM instruments (Bailey et al., 2022; Davis, 1989) and IDLE scales (Lee & Drajati, 2019; Lee et al., 2021). The results revealed that the respondents generally expressed that they used ChatGPT for various IDLE purposes. The authors also found that participants' positive attitude strongly predicted their intention of using the tool in the future, and this intention significantly affected their actual use. Liu and Ma's (2023) study has greatly contributed to our understanding of learner perceptions vis-à-vis ChatGPT use; however, the outcomes of their research is hardly generalizable to other settings as they used purposive sampling and recruited participants from several social media platforms, which may lead to selection bias. Furthermore, the study lacks additional external variables that might have an impact on participants' ChatGPT acceptance and use. Similar results were reached by a qualitative investigation conducted by Xiao and Zhi (2023) in China. The subjects were five graduate learners from various departments. Findings indicated that the participants had a positive attitude towards using ChatGPT for L2 learning, and they deemed ChatGPT as a personal tutor that delivers tailored feedback. This research can only provide initial insights about students' opinions on ChatGPT use since it was a small-scale study. Contrasting outcomes were reached by Yan (2023), who set out to explore how Chinese EFL learners perceive using ChatGPT in L2 writing. The findings indicated that students were aware of the benefits of this cutting-edge technology, yet they had concerns related to academic integrity. This study also used semi-structured interviews to gain insights from only eight students. Another research with the same purpose as the pre-stated studies was carried out by Singh et al. (2023) in a different setting, the UK. The researchers sampled 430 native and non-native English-speaking students from the department of Computer Science and used a questionnaire to capture their opinions about ChatGPT use for academic purposes. The results suggested that a great majority of the students were not using the tool frequently for academic purposes, and they expressed doubts about its potential benefits for learning. Being limited to only students from the department of Computer Science makes it hard to generalize the findings of this study. In another small-scale research carried out in the United Arab Emirates, Shoufan (2023) interviewed 56 senior students enrolled in the Computer Engineering program to develop a survey that explored learners' perceptions of using ChatGPT for learning. The most obvious finding to emerge from this study is that students had positive attitudes towards the tool and found it interesting and easy to use. However, they were divided in terms of the challenges it poses for learning. Adopting the unified theory of acceptance and use of technology two (UTAUT2) model, Romero-Rodríguez et al.'s (2023) relatively large-scale study surveyed 400 Spanish university students and found that gender was not an effective variable in determining ChatGPT use, but experience of using ChatGPT significantly impacted the facilitating conditions of the tool. Caution must be taken while interpreting the results of this study due to "the difficulty of inferring causality and the inability to assess changes over time" (Romero-Rodríguez et al., 2023, p. 336). In view of these limitations and contradictory outcomes, our study aims to address the stated gaps and broaden current knowledge of ChatGPT use of EFL learners by using a large

number of respondents who study at various English majors across different cohorts in different research context and gauging whether personality, as a psychological factor, impacts students' leveraging ChatGPT for LBC.

The Technology Acceptance Model

Among the various frameworks that have been developed for examining how people accept and use new technologies, Davis's (1989) technology acceptance model (TAM) stands out since it is widely acknowledged as a significant model that accurately predicts users' adoption of new technologies (Bailey et al., 2022). Figure 1 lists the components of the TAM (Liu & Ma, 2023, p. 4).

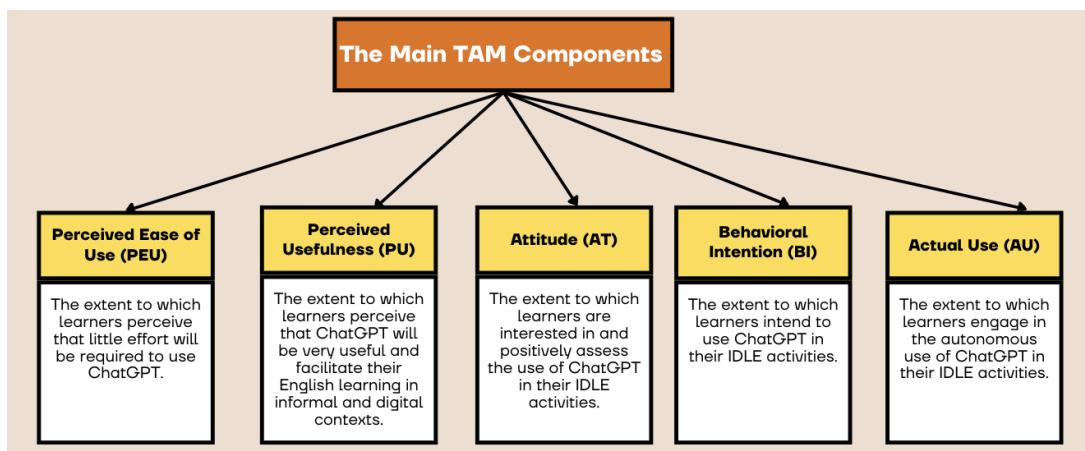


Figure 1. The main TAM components

Note. From "Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model" by G. Liu and C. Ma, 2023, *Innovation in Language Learning and Teaching*, p. 4. Copyright 2024 by Informa UK Limited.

According to Davis (1989), perceived ease of use (PEU) and perceived usefulness (PU) are two crucial factors that determine one's attitude (AT), which can influence individuals' behavioral intention (BI) to use a certain technology; this intention impacts the actual use (AU) of this technology to accomplish particular tasks. This study uses the TAM to investigate how EFL learners perceive, accept, and use ChatGPT in IDLE as empirical evidence confirmed the efficacy of the TAM in predicting the acceptance, rejection, or continuity of utilizing new technology (Faruk et al., 2023). The integration of the TAM into our research not only strengthens its theoretical underpinnings but provides a comprehensive view of the convoluted dynamics between EFL students' perceptions and their use of ChatGPT outside the classroom.

Personality and ChatGPT Use

Personality refers to "the dynamic organization within the individual of those psychological systems determining his unique adjustment for environment" (Allport, 1937, p. 48). Type-based and trait-based theories are the two unique approaches that examine personality. The type-based view has been challenged for being a static view that falls short of explaining the complicated nature of human personality. Accordingly, trait-based theories are supported by various researchers because "traits can be studied with a wide range of statistical techniques - all those methods based on the assumption of a normal distribution and continuous scores (means, standard deviations, etc.)" (Quenk, 1993, p. 9).

This study uses the *Five-Factor Model* (FFM) as a theoretical framework, which is currently the most widely recognized personality model and the most accurate estimate of the fundamental dimensions of

personality (McCrae & Costa, 2003). This model, commonly referred to as *The Big Five Model* of personality (BFM), has addressed the criticisms levelled against the earlier trait-based theories. *Openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism* are the five primary dimensions in this model (McCrae & Costa, 2003). Openness is associated with imagination, curiosity and creativity. High scorers on conscientiousness are hardworking, punctual, and organized, while those who score high in the extraversion trait tend to be social, talkative, and outgoing. Agreeableness relates to social harmony and cooperativeness. Lastly, neuroticism is characterized by anger, impulsiveness and depression. According to the BF model, each of the five dimensions is represented by six fundamental facets that each correspond to a particular domain.

Personality almost certainly determines human behavior (Saberi et al., 2021); therefore, from a technological perspective, the question is whether it affects people's use of ChatGPT. There has been a bulk of empirical investigations on the connection between personality and technology acceptance in general (Oksanen et al., 2020); however, they yielded varying outcomes. Oksanen et al. (2020) found positive correlations between openness and AI acceptance. Nevertheless, Svendsen et al. (2013) indicated that the intention to adopt new software was strongly associated with extraversion and agreeableness dimensions. In their study using 180 freshman students, Devaraj et al. (2008) aimed to reveal the potential relationship between personality and the TAM regarding the use of collaborative technologies for education. The researchers found various significant correlations between each personality trait and subdimensions of the TAM. Rivers (2021) conducted a study with 149 Japanese university students and found that agreeableness and conscientiousness traits have an impact on actual Moodle use. In each case, personality profiles of the participants were found to significantly shape their technology acceptance and use. None of these studies focused specifically on ChatGPT, though. Accordingly, this research seeks to investigate the possible relationship between the personality traits of EFL learners and their use of ChatGPT in out-of-classroom settings in an attempt to bridge this research gap and to move this line of inquiry forward. Drawing on the results of prior research, it seems reasonable to anticipate that personality might affect EFL students' ChatGPT use and perceptions. Indeed, very few studies delved into the association between personality and ChatGPT use. To illustrate, Li (2023) sampled 943 Chinese respondents from various backgrounds and found that openness mediated the impact of intellectual humility on embracing ChatGPT. Li's (2023) research is noteworthy with its large sample size and sound methodology, yet it is only concerned with openness, neglecting other personality traits. De Winter et al.'s (2023) study deserves particular attention for being a large-scale investigation into the relationship between individuals' ChatGPT use and their personality profiles. The researchers conducted their study in 34 countries with 2000 respondents who were either students or full-time employees. They sought to explore the frequency and acceptance of ChatGPT use and found that 32% of the subjects were using the tool on a weekly basis. The results also demonstrated that participants valued the benefits of ChatGPT although they expressed moral concerns. Respondents with high neuroticism levels were found to use the tool less frequently. As with any research, De Winter et al.'s (2023) study is not free from limitations. One major drawback lies in its reliance on the prolific population that risks self-selection and demographic biases which may affect the generalizability of the results, as acknowledged by the researchers themselves. In addition, methodological transparency was disregarded by the authors since they did not give much detail about the method, sampling technique or the personality scale used in their research. Li's (2023) and De Winter et al.'s (2023) studies focused on using ChatGPT for various purposes, not for educational use. To fill that void, Faruk et al. (2023) adopted a quantitative survey design and sampled 177 students from three Thai and Indian universities. The researchers revealed that openness had a favorable effect on ChatGPT use, while neuroticism and agreeableness had negative effects. Despite delving into an under-represented topic, Faruk et al.'s (2023) research suffers from small sample size, and it only demonstrates the issue from the eyes of Asian students, concentrating only on the use of ChatGPT for general educational purposes. Our study is more focused in this regard as it is about using ChatGPT for L2 learning and has a larger sample size.

It is apparent that ChatGPT has revolutionized the way we teach and learn. Thus, capturing users' intentions, perceptions and trust in this ubiquitous tool and unveiling the factors that might impact these perceptions and attitudes are important. In light of this, the present research offers insightful information since it unveils EFL students' frequency and purposes for using ChatGPT in extramural IDLE settings.

Research questions

Previous research has not adequately delved into the interplay between EFL students' ChatGPT use and their personality profiles. Given the significance of different personality factors in L2 acquisition and technology acceptance, the present study seeks to address this gap by examining the possible association between these two variables. In light of this niche and to achieve the pre-stated research goals, this study is guided by the following research questions (RQs):

- RQ 1. What are the levels of personality traits of Turkish EFL learners?
- RQ 2. How often do EFL learners use ChatGPT for IDLE?
- RQ 3. How do EFL learners perceive and use ChatGPT for IDLE?
- RQ 4. Is there a relationship between the personality traits of EFL learners and their frequency of using ChatGPT for IDLE?
- RQ 5. Is there a relationship between the personality traits of EFL learners and their perceptions of using ChatGPT for IDLE?

Method

To investigate EFL learners' acceptance and use of ChatGPT in IDLE, the present study adopted the non-experimental *quantitative methodology*, which draws on deductive reasoning and positivist paradigm (Creswell & Creswell, 2018). This research method was a good match with our research problem and objectives. The *cross-sectional survey design* was used to collect data since it is practical and effective in obtaining data from a large sample that would be impractical to assess directly.

Participants

This study was carried out at the English major departments of a public university in the eastern part of Türkiye during the fall semester of the 2023-2024 academic year. This particular location was chosen because both of the researchers work as full-time lecturers at the School of Foreign Languages at this university, making it easy to access the participants. Accordingly, convenience sampling was used in the data gathering procedure. Data regarding the demographic characteristics of the respondents, who were 535 native Turkish speaker undergraduates, were demonstrated in greater detail in Figure 2.

	<i>n</i>	%
Gender		
Female	351	65.6
Male	179	33.5
Prefer not to say	5	0.9
Age		
18-30	526	98.3
31-31+	9	1.7
Department		
American Culture and Literature	57	10.7
English Language and Literature	112	20.9
English Language Teaching	212	39.6
Translation and Interpreting	154	28.8
Grade		
Preparatory Class	126	23.6
Freshman	190	35.5
Sophomore	82	15.3
Junior	115	21.5
Senior	22	4.1
Perceived Level of English Proficiency		
A1	4	0.7
A2	17	3.2
B1	176	32.9
B2	248	46.4
C1	84	15.7
C2	6	1.1
Total	535	100

Figure 2. Demographic characteristics of the participants

13.1. Instrument(s)

Quantitative data were gathered by using three tools: The Demographic Information Form, The Big Five Inventory (BFI) and the TAM Questionnaire. Before commencing data-collection, all the necessary permissions were granted from the developers of these tools, and two seasoned scholars from the departments English Language Teaching as well as Psychological Counselling and Guidance reviewed them for validity.

2.2.1. The demographic information form

At the beginning of the data gathering process, the participants were asked to complete a Demographic Information Form consisting of five questions regarding gender, age, department, grade as well as perceived English proficiency level (Figure 2).

2.2.2. The Big Five inventory (BFI)

Once the students completed the demographic information form, they filled out the 44-item Big Five Inventory (Appendix A). Being one of the most commonly utilized instruments to assess personality profiles, this scale is based on a 5-point Likert scale (1 = disagree strongly; 5 = agree strongly). To make sure that it was simple enough for our sample to understand, we leveraged the Turkish version of the scale that was adapted and validated by Karaman et al. (2010). In their study, Karaman et al. (2010) used a sample of 1153 undergraduates to examine the validity and reliability of their scale and reported satisfactory levels of Cronbach's alpha values ranging from .75 to .86 for its subscales. The reliability measures in our study indicated a value of .72 for the overall scale, and values spanning from .71 to .75 for the subscales, which are all acceptable ranges.

2.2.3. The TAM questionnaire

The 22-item TAM questionnaire (Appendix B), created by Liu and Ma (2023) was administered later to gather data on participants perceptions and use of ChatGPT for English learning outside the classroom. The tool is rated on a six-point Likert scale with 1 used for ‘strongly disagree’ and 6 for ‘strongly agree’. It includes the five TAM subscales. In this study, the original English version was utilized since the tool was simple enough to be understood by even low-proficiency learners, and our sample students were proficient in both languages. In our study, the reliability score for the entire questionnaire was .95, and between .75 and .89 for the sub-components, showing satisfactory levels of reliability. We asked the frequency of the participants’ ChatGPT use for IDLE before the first question of the TAM questionnaire. Those who responded as ‘never’ were directed to the end of the section and were not permitted to complete the TAM questionnaire.

13.2. Data collection procedures

All the necessary legal and ethical procedures were fulfilled prior to the collection of data (the institution where this study was conducted, Social and Human Sciences Ethics Committee, approval number: E.88656144-000-2400001836). The subjects were informed about the voluntary nature of participation and guaranteed confidentiality. After the purpose of the study was explained, the data collection tools were administered in supervision of the researchers during class time and via a link to Google Forms to make it impossible to skip a question and avoid missing data.

13.3. Data analysis

The Statistical Package for Social Sciences (SPSS, version 20) program was used to analyse the data collected from 535 participants, administering both descriptive and inferential techniques. Descriptive statistical analyses were used to answer the first three RQs, to reveal the sample characteristics, their frequency, and perceptions of ChatGPT for IDLE. A Pearson product-moment correlation coefficient was computed for RQs four and five.

14. Results

14.1. Students’ Personality Traits

Table 1 displays the levels of personality traits of the participants according to subscales of the BFI, as asked by the first RQ.

Table 1. Descriptive statistics results for the participants’ personality traits

Personality traits	<i>M</i>	<i>SD</i>
Extraversion	3.32	.40
Agreeableness	3.40	.36
Conscientiousness	3.61	.36
Neuroticism	3.22	.41
Openness to Experience	3.68	.44

N= 535

The learners’ personality characteristics matched the *Openness to Experience* subscale to the highest extent ($M = 3.68$, $SD = 0.44$). Their least prevalent personality dimension was *Neuroticism* ($M = 3.22$, $SD = 0.41$).

Frequency of ChatGPT Use

According to the results, the participants' frequency of ChatGPT use was at a moderate level ($M = 2.56$), showing a considerable amount of diversity ($SD = 1.51$). 192 (35.9 %) of the respondents have never used ChatGPT for their IDLE. 108 (20.2 %) used the tool once every few months; 75 (14.0 %) and 64 (12.0 %) of the subjects expressed that they use ChatGPT once a month and once a week, respectively. Finally, 96 (17.9 %) of them use it several times a week. Therefore, it was revealed that the majority of the participants used ChatGPT for their IDLE; however, since the frequency of their usage was not so high, they may not be deemed as regular users.

Learners' Perceptions and Use of ChatGPT

Learners' perceptions and use of ChatGPT for IDLE were presented in Table 2.

Table 2. Descriptive statistics results for learners' perceptions and use of ChatGPT

	<i>M</i>	<i>SD</i>
Learners' Overall Perception and Use of ChatGPT	4.59	.84
Perceived Ease of Use	4.85	.83
Perceived Usefulness	4.71	.92
Attitude	4.68	.93
Behavioral Intention	4.42	1.04
Actual Use	4.32	1.04

Learners' perceptions and use of ChatGPT for their IDLE were quite positive and high ($M = 4.59$). The low standard deviation ($SD = 0.84$) indicated that there was a general agreement about overall perceptions and usage among the subjects. When it comes to how participants perceive ChatGPT and for which feature they use it the most, *PEU* was particularly noteworthy ($M = 4.85$, $SD = 0.83$). The one that the participants considered least important was *AU* ($M = 4.32$, $SD = 1.04$). However, when all subscales are examined, it is seen that the mean values of them were all high. Therefore, although the participants appreciated different features of ChatGPT and preferred it due to its different features, it is clear that they found this AI-powered tool useful and included it in their self-directed informal learning of English. The relatively high standard deviations for *BI* and *AU* ($SD = 1.04$) suggests that there were certain variations among users since some of them may plan to use the tool in the future more than others or some learners may use ChatGPT more extensively than the other students.

The Relationship Between the Personality Traits and Frequency of ChatGPT Use

The fourth RQ investigated whether there was a relationship between EFL learners' personality traits and their frequency of ChatGPT use, and the results are demonstrated in Table 3.

Table 3. Correlation Results for the Personality Traits and Frequency of ChatGPT Use

	1	2	3	4	5	6
Frequency of ChatGPT Use	1					
Extraversion	.13**	1				
Agreeableness	.09*	.26**	1			
Conscientiousness	.07	.31**	.31**	1		
Neuroticism	-.04	.03	.32**	.25**	1	
Openness to Experience	.10*	.43**	.26**	.29**	.08	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

As Table 3 demonstrates, there was a small and positive correlation between *Extraversion* and the frequency of using ChatGPT ($r = .13$), which was statistically significant ($p = .01$). Same applies for *Agreeableness* ($r = .09$), which was statistically significant ($p = .05$). There was a small and positive correlation between *Openness to experience* and the frequency of ChatGPT use ($r = .10$), which was statistically significant ($p = .05$), as well. All in all, although there is a significant relationship between some personality traits and the frequency of ChatGPT use, it is difficult to talk about a very strong relationship between them. Nevertheless, it can be stated that students who are extraverted, agreeable, and open to new experiences tended to use ChatGPT more frequently.

The Relationship Between Personality Traits and Perceptions of Using ChatGPT

Table 4 shows the results related to the relationship between personality and ChatGPT use.

Table 4. Correlation results for the personality traits and perceptions of ChatGPT use

	1	2	3	4	5	6	7	8	9	10	11	12
Personality traits	1											
TAM	.19**	1										
Extraversion	.64**	.20**	1									
Agreeableness	.65**	.17**	.26**	1								
Conscientiousness	.66**	.22**	.31**	.31**	1							
Neuroticism	.51**	-.06	.03	.32**	.25**	1						
Openness to Experience	.71**	.09	.43**	.26**	.30**	.08	1					
Perceived Ease of Use	.11*	.79**	.08	.10	.17**	.02	.02	1				
Perceived Usefulness	.18**	.92**	.19**	.17**	.22**	-.04	.07	.69**	1			
Attitude	.15**	.90**	.17**	.13*	.18**	-.08	.09	.69**	.81**	1		
Behavioral Intention	.16**	.82**	.17**	.15**	.22**	-.06	.06	.57**	.70**	.68**	1	
Actual Use	.19**	.91**	.22**	.18**	.18**	-.08	.12*	.58**	.78**	.75**	.69**	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

It can be understood from the table that there was a statistically significant correlation between the participants' personality traits and their perceptions regarding ChatGPT use. A small and positive correlation was observed between *Extraversion* and *AU* ($r = .22$), *Conscientiousness* and *PU* ($r = .22$), and *Conscientiousness* and *BI* ($r = .22$), which were all statistically significant ($p = .01$). Same applied for *Extraversion* and *PU* ($r = .19$), which was statistically significant ($p = .01$). Likewise, *Agreeableness* and *AU* ($r = .18$), *Conscientiousness* and *AT* ($r = .18$), as well as *Conscientiousness* and *AU* ($r = .22$) showed statistically significant positive correlations ($p = .01$). A statistically significant ($p = .01$) positive correlation was found between *Extraversion* and *AT* ($r = .17$), *Extraversion* and *BI* ($r = .17$), *Agreeableness* and *PU* ($r = .17$), and *Conscientiousness* and *PEU* ($r = .17$). *Agreeableness* and *BI* ($r = .15$) were positively correlated, as well. Finally, *Agreeableness* and *AT* ($r = .13$) and *Openness to experience* and *AU* ($r = .12$) showed positive correlations that were statistically significant ($p = .05$). Overall, although the correlation was small, there was a statistically significant association between some personality dimensions and perceptions regarding ChatGPT use.

15. Discussion

Students' personality traits

The purpose of the first RQ was to investigate the personality traits of the participants. Learners were found to be highly open to experience, suggesting that students are open to trying new approaches while learning a foreign language. However, in an empirical investigation that was carried out by Asmali in 2014, different outcomes emerged although the researcher conducted the study in Turkish EFL context, as well. According to Asmali (2014), the most dominant personality dimension was agreeableness. Such a contrasting result might be explained by the specific sample that was chosen for each study. Unlike the present research, Asmali (2014) recruited elementary-level learners enrolled in various departments of a public university in Turkey. In addition, the conscientiousness characteristic of our participants was also quite high, which may suggest that they act with a sense of duty and that if they are willing to learn a foreign language, they will be open and willing to benefit from any technological tool they perceive useful. Likewise, Cachero et al. (2023) express that as conscientious learners are prudent and responsible, they demonstrate high levels of performance and a strong sense of ambition and determination. Faruk et al. (2023) adopted a similar research design to our study and found that openness to experience had a positive effect on ChatGPT use. Therefore, it seems reasonable to assume that this finding is promising in terms of ChatGPT use since learners with this personality trait are likely to choose to leverage ChatGPT for their IDLE.

Frequency of ChatGPT use

With respect to the second RQ, it was disclosed that 343 out of the 535 respondents have used ChatGPT for IDLE before, which accounts for more than sixty percent. Considering that a very short time has passed since ChatGPT entered our lives, this rate can be considered quite appreciable. However, the findings of our study are contradictory with those of Singh et al.'s (2023) research conducted on the frequency of ChatGPT use in a different context. The researchers found that 55 % of the subjects have never utilized ChatGPT. The reason for this difference may be that their study was on students' use of ChatGPT for academic purposes, while ours was about the use of ChatGPT for learning English in informal digital contexts, in particular. Although concerns about the reliability of ChatGPT negatively affect its use in academic terms, it may be preferred for out-of-class learning due to factors such as ease of use and practicality. Studying at an English major department might also be a predicting factor regarding ChatGPT use, too. Furthermore, as put by Bin-Hady et al. (2023), ChatGPT is a user-friendly tool for language learning; it may have been preferred by our participants because it was perceived as such in our context. This implies that EFL learners have adopted ChatGPT in terms of IDLE, which is a promising finding. However, taking into consideration that 20.2 % of the participants use ChatGPT once every few months, we can conclude that it is not fully embraced by EFL learners, which means that students do not use ChatGPT quite often for learning English on their own. This result is consistent with the outcomes of De Winter et al.'s (2023) study, which found that 20% of the respondents used the tool less than once a month. Such an infrequent engagement with the tool might suggest the availability and usage of some other AI technologies, students' hesitations regarding academic integrity and the accuracy of the content that ChatGPT produces, or the policies adopted by the institution in which they study about not using any form of AI tool for their assignments. Accordingly, studies which use student interviews to dig into the reasons for not using ChatGPT that often are warranted. In addition, the outcomes related to ChatGPT usage frequency also

suggested that a considerable amount of diversity exists among the users ($SD = 1.51$); thus, factors which might impact different usage patterns should be investigated by further studies.

Learners' perceptions and use of ChatGPT for IDLE

Learners' perceptions of ChatGPT use were positive, and a sizable number of them used this cutting-edge tool for IDLE. It is obvious that the results of our research align with the study of Xiao and Zhi (2023). Likewise, researching tertiary level students' readiness and PU for using ChatGPT for academic purposes, Adams et al. (2023) found that students had a positive perception towards the impact of ChatGPT on their learning. The positive perception and widespread use of a new tool in such a short time can be explained by its PEU, which is the reason for the preference of most of the participants. Considering that university students take many courses at the same time and race against time to complete many assignments on time, they may not be able to devote much time for their IDLE. Thus, ease of use enables students to save both effort and time in this respect. In a similar context, preservice primary school teachers' perceptions about ChatGPT were investigated, and 87% of the participants stated that they perceived ChatGPT as easy to access and use (Lozano & Fontao, 2023). Furthermore, our participants also attributed value to ChatGPT in terms of PU, AT, BI, and AU, and consider it as an important component for their IDLE. Liu and Ma (2023) also stated that they found AU as an indicator of their learners' use of ChatGPT for their English language learning activities beyond the classroom. From this perspective, we can conclude that the participants of this study also exhibited positive attitudes towards autonomous use of ChatGPT in their IDLE.

The relationship between students' personality traits and frequency of ChatGPT use

One of the preliminary works on the personality traits of learners in terms of technology acceptance and use was undertaken by Devaraj et al. (2008). They found that people with high levels of conscientiousness, extraversion, agreeableness, and neuroticism intend to use new technology and adopt its AU. We found a statistically significant but low-impact relationship in terms of extraversion, agreeableness, and openness to experience. At this point, our results match those of Devaraj et al. (2008) only in terms of extraversion and agreeableness. The outcome that neuroticism did not correlate with the use of ChatGPT, which was expected, can be explained by the disposition of these learners to be anxious because individuals with this personality trait are generally not very open to innovations, and they do not like to be up in the air. In a similar vein, De Winter et al.'s (2023) study revealed that learners who scored high in neuroticism were the group that used ChatGPT the least. Moreover, the reason why there is no significant relationship between having the trait of conscientiousness and the frequency of using ChatGPT can be explained by the notion that students with this personality trait may have been affected by the discussions about whether the use of ChatGPT is ethical or not and whether its sources are reliable or not. Since these people think in great detail and handle everything meticulously, they may not trust a tool whose accuracy and reliability are questioned. However, Faruk et al. (2023) state that ChatGPT's ability to provide adaptable and personalized learning experiences is an important element in expanding its use for educational purposes. Therefore, as the demerits regarding ChatGPT are eliminated and evidence-based studies on its usefulness increase, students' perspectives will move towards a positive direction, so will their frequency of utilizing the tool.

The relationship between students' personality traits and their perceptions of ChatGPT use

Regarding the final RQ, significant, though not large, relationships emerged between four of the personality traits and perceptions toward ChatGPT use for IDLE. In De Winter et al.'s (2023) study, the strength of the correlations between ChatGPT use and personality traits was also weak. We can attribute this to the generic and context-independent nature of personality traits. In our study, the only personality trait that showed no significant relationship with ChatGPT use and perception was neuroticism, a finding that is also consistent with the results of De Winter et al. (2023). They also noted that neurotic individuals tended to be more anxious about using ChatGPT. Thus, it can be concluded that due to the fluctuating emotions and anxiety-prone personality features of neurotic people, participants may have developed suspicious perceptions toward ChatGPT use. Devaraj et al. (2008) stated that the big five personality dimensions were related to the basic dimensions of technology acceptance, and our study overlaps with their study in terms of ChatGPT use and perception when neuroticism is excluded. Therefore, it would be a more efficient approach to evaluate other personality traits separately in terms of ChatGPT use and perception.

To begin with, extroverted individuals are open to new information and communication, and they are known to be willing to embark on novel adventures. For this reason, these students may tend to use ChatGPT for its PU, AT, BI, and AU. Extroverted students may have the intention to adopt a new software (Svendsen et al., 2013). Thus, these learners may have easily adopted a new and practical language learning tool for IDLE due to its PU. Moreover, it is clear that these students have a positive attitude towards using ChatGPT to conduct IDLE activities, and they tend to use it autonomously for different English learning activities in line with the AU of ChatGPT.

Moreover, in this study, a statistically significant relationship emerged between being conscientious and the subscales of PU and BI. These results are supported by Devaraj et al.'s (2008) study. As conscientious people have high self-discipline and sense of duty, they are often willing to adopt what they believe is useful to achieve their goals; therefore, it makes sense for conscientious EFL learners to use ChatGPT as a source in their IDLE due to its PU and BI. There is also a significant relationship, albeit on a small scale, between conscientiousness and PEU, AT, and AU.

Regarding agreeableness, a statistically significant but small-scale relationship was found between this trait and AU, PU, BI and AT. Devaraj et al. (2008) also suggested a notable relationship between agreeableness and PU. Agreeable people are often solution-oriented and look for the best and healthiest way to accomplish tasks. These people, who tend to cooperate and work in harmony with others, may feel lonely when it comes to IDLE because they are used to completing tasks with their peers in the classroom environment and learning by supporting each other. Accordingly, they may see ChatGPT as their peers with whom they can work together and complete their IDLE activities. Furthermore, Bin-Hady et al. (2023) state that ChatGPT creates fun learning environments and offers custom tailored game-like activities for learners. Hence, participants may have perceived ChatGPT as useful and developed a positive attitude toward it. This may be the reason why they have BI towards using it and a positive understanding towards its AU.

Lastly, openness to experience showed a statistically significant, albeit small, relationship only with AU, which is quite surprising, considering the characteristics of this personality trait, such as creativity, curiosity, and supporting original ideas. Our results are incompatible with those of Ozbek et al. (2014), who found a significant relationship between openness and PEU in their study. However, it

can be concluded that their disposition for intellectual pursuits may have led them to keep their distance to ChatGPT. For this reason, they may want to construct knowledge based on their own intellectual endowment. Overall, this study provides valuable insights in terms of EFL learners' perceptions of ChatGPT use for IDLE; however, more research is warranted to generalize its findings to other educational settings.

16. Conclusions

This research seeks to unveil EFL students' personality profiles, examine their frequency and perceptions of using ChatGPT for learning English outside the conventional classroom settings and scrutinizing a hitherto neglected field by digging into the possible connection between these variables. By deepening our knowledge of ChatGPT use for students, we can have a potential impact on further research, policy, and practice.

Pedagogical Implications

In light of the findings, two major implications were obtained. First of all, EFL learners have already started using ChatGPT for IDLE, and they have positive perceptions towards its use. Therefore, instructors should integrate this tool in formal academic environments by using blended learning activities. Secondly, given that a significant correlation was found between some personality traits and ChatGPT perception, although the effect rates were small scale, students with different personality traits can use ChatGPT in their L2 learning, at least as a facilitator. Instructors should use various AI technologies to encourage their learners' autonomous language learning beyond the conventional classroom walls, by designing more tailored tasks to meet their personality-based needs.

Limitations of the Study

It is important to note a few limitations of this study. Since this research is a cross-sectional study, its results convey preliminary data, which does not allow causal inferences. By conducting a comprehensive and well-planned mixed methods study, more insightful results can be obtained. Furthermore, since the results were based on self-report data, students may not have been able to clearly express their heartfelt thoughts. Therefore, the possibility that they shared opinions with higher social acceptance should also be taken into consideration. Finally, these data provide information about the context in which they occur and cannot be generalized to the entire EFL context. Despite the caveats of this research, significant contributions were achieved to enhance our knowledge regarding EFL learners' ChatGPT use and personality profiles.

Suggestions for Further Research

Although the results do not indicate causality, unearthing the relationship between personality traits and ChatGPT use has revealed whether students with variable personality traits tend to benefit from ChatGPT in IDLE. The differences found between the results of this study and those of other studies suggest that personality dimensions play a discriminating role on BI regarding the use of ChatGPT in the IDLE setting. Therefore, it would be useful to reconsider and repeat this study in different educational contexts. It is also advisable to adopt mixed methods approaches to determine if age, culture, proficiency level, aptitude, motivation, and educational context are well-grounded variables

influencing students' perceptions and frequency of ChatGPT use. They could be the mediating or moderating factors that may affect the observed relationship between personality and ChatGPT use, as well. Moreover, as the voice of all stakeholders is important for education-related decisions, it is of great importance to investigate teachers' opinions and preferences regarding the use of ChatGPT for learning English.

References

Adams, D., Chuah, K., Devadason, E., & Azzis, M. S. A. (2023). From novice to navigator: Students' academic help-seeking behaviour, readiness, and perceived usefulness of ChatGPT in learning. *Education and Information Technologies*, 29(11), 13617–13634. <https://doi.org/10.1007/s10639-023-12427-8>

Allport, G. W. (1937). *Personality: A psychological interpretation*. Henry Holt.

Asmali, M. (2014). The relationship between the Big Five personality traits and language learning strategies. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 17(32), 1-18.

Bailey, D. R., Almusharraf, N., & Almusharraf, A. (2022). Video conferencing in the e-learning context: Explaining learning outcome with the technology acceptance model. *Education and Information Technologies*, 27(6), 7679-7698. <https://doi.org/10.1007/s10639-022-10949-1>

Barnett, T., Pearson, A. W., Pearson, R., & Kellermanns, F. W. (2015). Five-factor model personality traits as predictors of perceived and actual usage of technology. *European Journal of Information Systems*, 24(4), 374-390. <https://doi.org/10.1057/ejis.2014.10>

Benson, P. (2011). Language learning and teaching beyond the classroom: An introduction to the field. In P. Benson, & H. Reinders (Eds.), *Beyond the language classroom* (pp. 7-16). Palgrave Macmillan.

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. <https://doi.org/10.1108/LHT-05-2023-0200>

Cachero, C., Rico-Juan, J. R., & Macià, H. (2023). Influence of personality and modality on peer assessment evaluation perceptions using machine learning techniques. *Expert Systems with Applications*, 213, 119150. <https://doi.org/10.1016/j.eswa.2022.119150>

ChatGPT sets record for fastest-growing user base - analyst note. (2023). Reuters. Available from <https://www.reuters.com/technology/chatgpt-sets-record-fastestgrowing-user-base-analyst-note-2023-02-01/>.

Chun, D. M. (2019). Current and future directions in TELL. *Journal of Educational Technology & Society*, 22(2), 14-25.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of Information Technology. *MIS Quarterly*, 13(3), 319. <https://doi.org/10.2307/249008>

Devaraj, S., Easley, R. F., & Crant, J. M. (2008). Research note—how does personality matter? Relating the five-factor model to technology acceptance and use. *Information Systems Research*, 19(1), 93–105. <https://doi.org/10.1287/isre.1070.0153>

De Winter, J., Dodou, D., & Eisma, Y. B. (2023). Personality and acceptance as predictors of ChatGPT use. *ResearchGate Preprints*.

Faruk, L. I., Rohan, R., Ninrutsirikun, U., & Pal, D. (2023). University students' acceptance and usage of generative AI (ChatGPT) from a psycho-technical perspective. *Proceedings of the 13th International Conference on Advances in Information Technology*. <https://doi.org/10.1145/3628454.3629552>

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 5a*. University of California, Berkeley, Institute of Personality and Social Research.

Karaman, N. G., Dogan, T., & Coban, A. E. (2010). A study to adapt the big five inventory to Turkish. *Procedia - Social and Behavioral Sciences*, 2(2), 2357–2359. <https://doi.org/10.1016/j.sbspro.2010.03.336>

Kovačević, D. (2023). Use of ChatGPT in ESP teaching process. In *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1-5). IEEE. <https://doi.org/10.1109/infoteh57020.2023.10094133>

Lee, J. S., & Drajati, N. A. (2019). Affective variables and informal digital learning of English: Keys to willingness to communicate in a second language. *Australasian Journal of Educational Technology*. <https://doi.org/10.14742/ajet.5177>

Lee, J. S., & Lee, K. (2021). The role of informal digital learning of English and L2 motivational self-system in foreign language enjoyment. *British Journal of Educational Technology*, 52(1), 358–373. <https://doi.org/10.1111/bjet.12955>

Lee, J. S., Xie, Q., & Lee, K. (2021). Informal digital learning of English and L2 willingness to communicate: Roles of emotions, gender, and educational stage. *Journal of Multilingual and Multicultural Development*, 1–17. <https://doi.org/10.1080/01434632.2021.1918699>

Li, H. (2023). Rethinking human excellence in the AI age: The relationship between intellectual humility and attitudes toward ChatGPT. *Personality and Individual Differences*, 215, 112401. <https://doi.org/10.1016/j.paid.2023.112401>

Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1–14. <https://doi.org/10.1080/17501229.2023.2240316>

Lozano, A., & Fontao, C. B. (2023). Is the education system prepared for the irruption of Artificial Intelligence? A study on the perceptions of students of primary education degree from a dual perspective: Current pupils and future teachers. *Education Sciences*, 13(7). <https://doi.org/10.3390/edusc13070733>

Ma, X., & Huo, Y. (2023). Are users willing to embrace ChatGPT? Exploring the factors on the acceptance of chatbots from the perspective of AIDUA framework. *Technology in Society*, 75. <https://doi.org/10.1016/j.techsoc.2023.102362>

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective*. Guilford Press.

Quenk, N. L. (1993). Personality types or personality traits: What difference does it make. *Bulletin of Psychological Type*, 16(2), 9-13.

Ozbek, V., Alniacik, U., Koc, F., Akkiliç, M. E., & Kas, E. (2014). The impact of personality on technology acceptance: A study on smart phone users. *Procedia—Social and Behavioral Sciences*, 150, 541–551. <https://doi.org/10.1016/j.sbspro.2014.09.073>

Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching*, 50(4), 561–578. <https://doi.org/10.1017/s0261444817000192>

Rivers, D. J. (2021). The role of personality traits and online academic self-efficacy in acceptance, actual use and achievement in Moodle. *Education and Information Technologies*, 26(4), 4353–4378. <https://doi.org/10.1007/s10639-021-10478-3>

Romero-Rodríguez, J.-M., Ramírez-Montoya, M.-S., Buenestado-Fernández, M., & Lara-Lara, F. (2023). Use of ChatGPT at university as a tool for complex thinking: Students' perceived usefulness. *Journal of New Approaches in Educational Research*, 12(2), 323. <https://doi.org/10.7821/naer.2023.7.1458>

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>

Saberi, M., DiPaola, S., & Bernardet, U. (2021). Expressing personality through non-verbal behaviour in real-time interaction. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.660895>

Shoufan, A. (2023). Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey. *IEEE Access*, 11, 38805–38818. <https://doi.org/10.1109/access.2023.3268224>

Singh, H., Tayarani-Najaran, M.-H., & Yaqoob, M. (2023). Exploring computer science students' perception of ChatGPT in higher education: A descriptive and correlation study. *Education Sciences*, 13(9). <https://doi.org/10.3390/educsci13090924>

Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B., & McLay, K. (2021). Informal Digital Learning of English (IDLE): A scoping review of what has been done and a look towards what is to come. *Computer Assisted Language Learning*, 36(4), 608–640. <https://doi.org/10.1080/09588221.2021.1936562>

Svendsen, G. B., Johnsen, J. K., Sorensen, L. A. and Vitterso, J. (2013). Personality and technology acceptance: The influence of personality factors on the core constructs of the technology acceptance model. *Behavior & Information Technology*, 32(4), 323–334. <https://doi.org/10.1080/0144929X.2011.553740>

Tiwari, C. K., Bhat, M. A., Khan, S. T., Subramaniam, R., & Khan, M. A. I. (2023). What drives students toward ChatGPT? An investigation of the factors influencing adoption and usage of ChatGPT. *Interactive Technology and Smart Education*. <https://doi.org/10.1108/itse-04-2023-0061>

Van Lier, L. (2000). From input to affordance: Social interactive learning from an ecological perspective. In J. P. Lantolf, (Ed.), *Sociocultural theory and second language learning: Recent advances* (pp. 245–259). OUP.

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.

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28(11), 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>

Appendix A. The Big Five inventory (John et al., 1991)

Disagree strongly Disagree a little Neither agree nor disagree Agree a little Agree strongly

1

2

3

4

5

I see myself as someone who ...

1. is talkative	1	2	3	4	5
2. tends to find fault with others	1	2	3	4	5
3. does a thorough job	1	2	3	4	5
4. is depressed, blue	1	2	3	4	5
5. is original, comes up with new ideas	1	2	3	4	5
6. is reserved	1	2	3	4	5
7. is helpful and unselfish with others	1	2	3	4	5
8. can be somewhat careless	1	2	3	4	5
9. is relaxed, handles stress well	1	2	3	4	5
10. is curious about many different things	1	2	3	4	5
11. is full of energy	1	2	3	4	5
12. starts quarrels with others	1	2	3	4	5
13. is a reliable worker	1	2	3	4	5
14. can be tense	1	2	3	4	5
15. is ingenious, a deep thinker	1	2	3	4	5
16. generates a lot of enthusiasm	1	2	3	4	5
17. has a forgiving nature	1	2	3	4	5
18. tends to be disorganized	1	2	3	4	5
19. worries a lot	1	2	3	4	5
20. has an active imagination	1	2	3	4	5
21. tends to be quiet	1	2	3	4	5
22. is generally trusting	1	2	3	4	5
23. tends to be lazy	1	2	3	4	5
24. is emotionally stable, not easily upset	1	2	3	4	5
25. is inventive	1	2	3	4	5
26. has an assertive personality	1	2	3	4	5
27. can be cold and aloof	1	2	3	4	5

28. perseveres until the task is finished	1	2	3	4	5
29. can be moody	1	2	3	4	5
30. values artistic, aesthetic experiences	1	2	3	4	5
31. is sometimes shy, inhibited	1	2	3	4	5
32. is considerate and kind to almost everyone	1	2	3	4	5
33. does things efficiently	1	2	3	4	5
34. remains calm in tense situations	1	2	3	4	5
35. prefers work that is routine	1	2	3	4	5
36. is outgoing, sociable	1	2	3	4	5
37. is sometimes rude to others	1	2	3	4	5
38. makes plans and follows through with them	1	2	3	4	5
39. gets nervous easily	1	2	3	4	5
40. likes to reflect, play with ideas	1	2	3	4	5
41. has few artistic interests	1	2	3	4	5
42. likes to cooperate with others	1	2	3	4	5
43. is easily distracted	1	2	3	4	5
44. is sophisticated in art, music, or literature	1	2	3	4	5

APPENDIX B. The 22-Item TAM questionnaire (Liu & Ma, 2023)

	① Strongly disagree	② Disagree	③ Slightly disagree	④ Slightly agree	⑤ Agree	⑥ Strongly agree
1. I believe that ChatGPT can help me improve the quality of my learning.	①	②	③	④	⑤	⑥
2. I find ChatGPT very convenient to use.	①	②	③	④	⑤	⑥
3. I think ChatGPT can help me learn more effectively.	①	②	③	④	⑤	⑥
4. I use ChatGPT to increase my interest and motivation in learning English beyond the classroom.	①	②	③	④	⑤	⑥
5. ChatGPT is a very attractive learning tool.	①	②	③	④	⑤	⑥
6. I use ChatGPT to improve my English communication skills beyond the classroom.	①	②	③	④	⑤	⑥
7. Learning with ChatGPT is very interesting.	①	②	③	④	⑤	⑥
8. I find it easy to understand the content provided by ChatGPT.	①	②	③	④	⑤	⑥
9. I am willing to spend time and effort to learn how to use ChatGPT better.	①	②	③	④	⑤	⑥
10. I expect to use ChatGPT frequently in the future for learning purposes.	①	②	③	④	⑤	⑥
11. I find ChatGPT useful in my extracurricular studies.	①	②	③	④	⑤	⑥
12. I use ChatGPT to enhance my English reading and writing abilities, such as proofreading in out-of-class settings.	①	②	③	④	⑤	⑥
13. I think ChatGPT can provide me with more learning opportunities.	①	②	③	④	⑤	⑥
14. I think ChatGPT can help me improve my learning abilities.	①	②	③	④	⑤	⑥
15. I enjoy learning with ChatGPT.	①	②	③	④	⑤	⑥
16. I plan to use ChatGPT for my learning needs in the future.	①	②	③	④	⑤	⑥
17. I think it is easy to learn how to use ChatGPT.	①	②	③	④	⑤	⑥
18. I feel that using ChatGPT is straightforward.	①	②	③	④	⑤	⑥
19. I use ChatGPT to efficiently gather and integrate English learning information in the informal and online settings.	①	②	③	④	⑤	⑥
20. I use ChatGPT to help me better understand the English language and culture in out-of-class settings.	①	②	③	④	⑤	⑥
21. Using ChatGPT for learning is a good idea.	①	②	③	④	⑤	⑥
22. I use ChatGPT to enhance my performance in English language learning beyond the classroom.	①	②	③	④	⑤	⑥

AUTHOR BIODATA

Zeynep Daşer is a lecturer at Ataturk University, School of Foreign Languages, Erzurum, Türkiye. She received her master's degree on English Education from Ufuk University, ELT Department. She is currently doing her PhD at Ataturk University, ELT Department. Her research interests include L2 writing, positive psychology, L2 writing assessment, written corrective feedback, digital multimodal composing, and using artificial intelligence in L2 learning.

Bahar Kar is a lecturer at Ataturk University, School of Foreign Languages, Erzurum, Türkiye. She earned her master's degree on English Education from Çağ University, ELT Department. She is currently pursuing her PhD at Ataturk University, ELT Department. Her research interests centre around learning and teaching English as a foreign language, positive psychology in language learning, and individual differences in foreign language education.

From Red Pen to Algorithm: A Systematic Review of AI and Teacher Feedback in Writing

Aysun Barış Kosaogulları ^{a 6}, Vasfiye Geçkin ^{b 2}

^a Girls High School & Democracy University, Izmir, Türkiye

^b Democracy University, Izmir, Türkiye

APA Citation:

Barış Kosaogulları, A., & Geçkin, V. (2025). From Red Pen to Algorithm: A Systematic Review of AI and Teacher Feedback in Writing

Abstract

Over the past few years, automated written feedback (AWF) has received increasing attention in language teaching and learning. This systematic review compares the feedback provided by teachers and AI tools, focusing on five key dimensions: research context, feedback type, learner response, effectiveness in improving writing, and pedagogical implications for classroom practice. While research context differs in learners' age and proficiency, most studies were implemented in higher education contexts. Among the eleven AI tools used in the studies, ChatGPT was the most frequently used one. AWF agents focused more on form, whereas human teachers concentrated on meaning and comprehension. Language learners reported that AWF was quick and accessible, but teacher feedback was more reliable and deeper. Hybrid feedback was found to be the most effective since it combines the higher-order improvements provided by teachers with the speed and form-level feedback delivered by AI tools. We suggest that authorities and policymakers integrate AI tools in feedback processes as a supporter, not a replacement for teachers, yet, to provide meaningful, accurate, and coherent feedback. Additionally, learners and teachers should have training in feedback literacy and the use of AI tools.

Keywords: automated writing feedback; computer writing feedback; computer essay feedback; second language writing; AI vs teacher feedback

Introduction

Feedback has a vital role in writing instruction, specifically in English language teaching (ELT) by providing guidance on how to improve writing skills, addressing areas such as content, organization, vocabulary, language use, or mechanics. Feedback has been defined by many scholars in the scope of writing instruction. Hyland and Hyland (2006) define feedback as "a key element in scaffolding students' writing development, providing guidance and support for improvement.". Likewise, Hattie and Timperley (2007) describe feedback as "information provided by an agent regarding aspects of one's performance or understanding, with the intention of improving learning.". They underpin the role of feedback affecting students' academic progress by accelerating self-regulation and iterative improvement. It is also underscored by Nicol and Macfarlane-Dick (2006) that feedback is a two-way, dialogic process instead of one-way information dissemination, emphasizing that students must react to and welcome feedback for it to be effective.

The nature of feedback in language teaching has changed recently with the rise of artificial intelligence (AI) tools such as Grammarly, Criterion, Write & Improve, and Pigai. Teachers were the only sources of feedback formerly, but now automated writing feedback (AWF) has gained a lot of attention since it enables immediate feedback and reduces the teacher workload (Han & Sari, 2024; Thi & Nikolov,

⁶ Corresponding author.

E-mail address: 2318116002@std.idu.edu.tr

2023). These AI tools primarily focus on form, leaving higher-order concerns to teachers (Yıldız & Kuru Gonen, 2024). However, with the release of Generative AI models such as ChatGPT and Gemini, providing feedback has expanded beyond error correction to include suggestions for organization, consistency, and style (Steiss et al., 2024; Zou et al., 2025). This transition triggered questions related to accuracy, reliability, and pedagogical implications of feedback provided by AI tools (Alsofyani & Barzanji, 2024; Tran, 2025).

Student and teacher perceptions are at the heart of this debate. While students generally appreciate the immediacy and accessibility of AI agents, they tend to trust and prefer teacher feedback, which is perceived as more reliable, supportive, and context-sensitive (Lo et al., 2025; Tran, 2025). On the other hand, teachers see potential in AWF since using AI reduces their burden, but there are still questions about its limitations, such as overgeneralizing or offering incidental incorrectness, and lacking insensitivity towards learner backgrounds (Guo & Wang, 2023). Combining both feedback types – depth of teacher feedback and AI's effectiveness – hybrid feedback models have emerged as the most auspicious way, facilitating accuracy, consistency, and student motivation (Sari & Han, 2024; Yıldız et al., 2025).

Although there is a growing body of research questioning the effectiveness of AI tools, comparative studies about AI and teacher feedback still cannot reach a consensus. In our systematic review, we compared the feedback provided by AI tools and teachers in five key dimensions: research context, feedback type, learner response, effectiveness in improving writing, and pedagogical implications for classroom practice. In this way, we hope to seek answers to the following research questions:

- 1) In what educational contexts have AI- and teacher-mediated feedback in L2 writing been investigated?
- 2) How do feedback types differ according to the feedback source in terms of focus and comprehensiveness?
- 3) How do learners respond to feedback provided by AI or teachers in terms of trust and preference?
- 4) What are the effects of AI vs. teacher feedback on improving L2 writing outcomes?
- 5) What pedagogical implications emerge from integrating AI and teacher feedback?

Methodology

2.1 Search Strategy

This review was carried out following the guidelines of systematic research (Preferred Reporting Items for Systematic Reviews and Meta-Analyses [PRISMA] 2020; Page et al., 2021). A comprehensive investigation was conducted across major academic databases, including Web of Science, Scopus, ERIC, Taylor & Francis Online, SpringerLink, and ScienceDirect between January 2022 and September 2025. Search keywords included terms related to feedback, such as: “AI feedback,

automated writing evaluation, teacher feedback, human feedback, second language writing, L2 writing, EFL, ESL".

The reference lists of the selected articles were also identified to determine whether there were any additional relevant studies. This preliminary search yielded 245 papers.

2.2 Inclusion & Exclusion Criteria

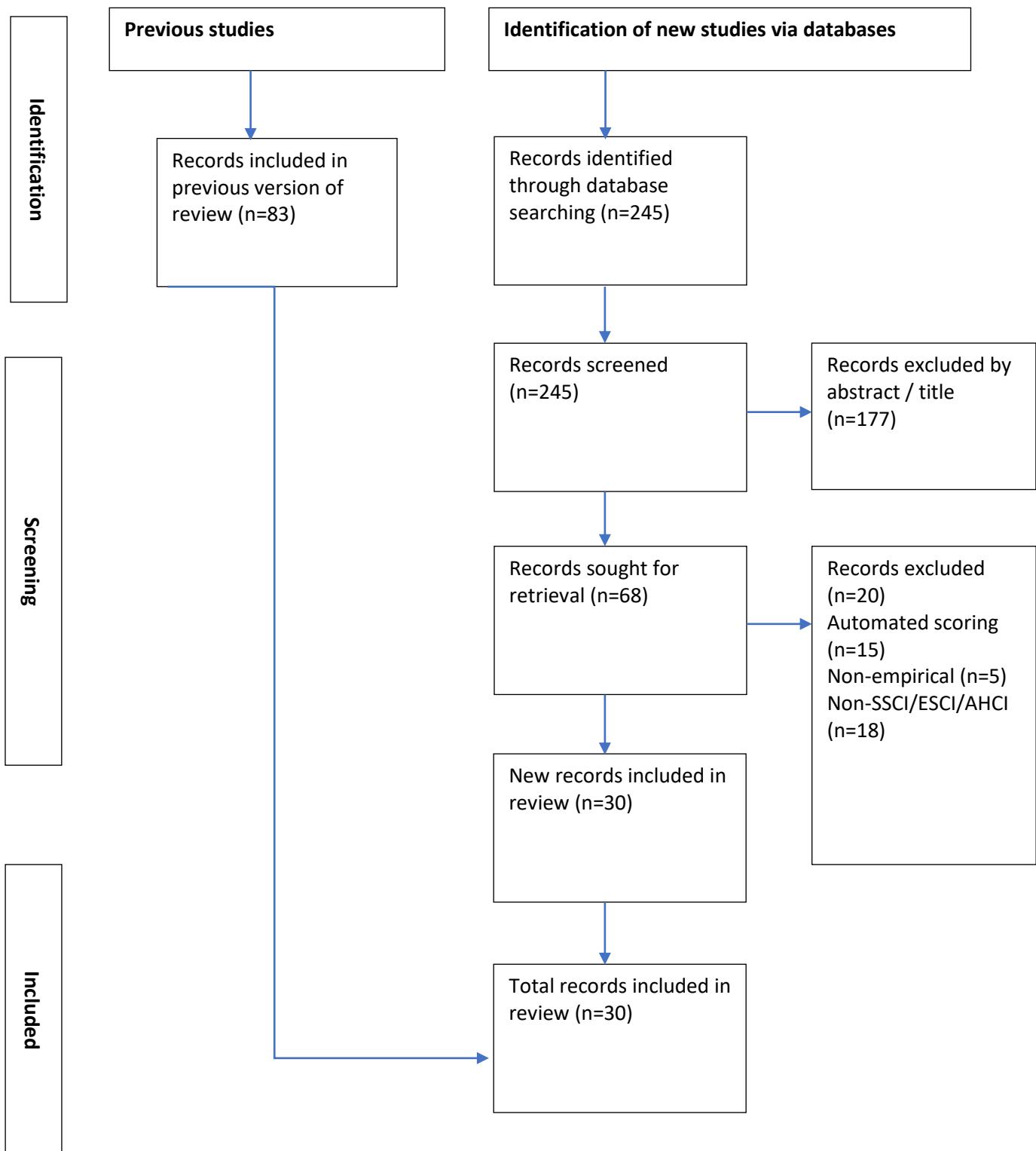
The first criterion was the relevance of the topic to the identified research. By reading the title and abstract, articles that did not compare AI feedback to teacher feedback in second language writing were excluded (n=177). Abstracts and research questions were reread to conduct a further screening, and studies focusing on automated writing scoring were excluded (n=15). The second criterion was whether the studies were experimental in nature, and non-experimental studies, such as review articles, were excluded (n=5). Our final criterion was that the articles must have been published in Arts & Humanities Citation Index (AHCI), Social Science Citation Index (SSCI), or Emerging Sources Citation Index (ESCI) journals, as these journals undergo rigorous peer review to ensure research quality (Duman, Orhon, & Gedik, 2015). Articles not listed in AHCI, SSCI, or ESCI were excluded (n=18). As a result, our study was finalized with a total of 30 articles. A PRISMA flow diagram (see Figure 1) illustrates the study selection process.

2.3 Data Extraction & Coding

To answer our research questions, a coding scheme was developed that included five analytical dimensions. The first category is the research context, which includes language proficiency level, language environment, and education level. The second category is feedback type, comparing AI-provided feedback to teacher-provided feedback in terms of focus (form vs meaning) and comprehensiveness. The third category is the learner response to the feedback mediated by AI tools or teachers in relation to students' trust and preferences. The next category is the effectiveness in improving writing accuracy, complexity, higher-order writing outcomes, and motivation. The final category is the pedagogical implications of the integration of AI and teacher feedback and hybrid models.

Figure 1.

PRISMA flow diagram representing the data selection process



Results & Discussion

3.1 Research Context

Table 1 presents the research context in comparative studies in relation to teacher vs AI feedback. The proficiency level of the students was mostly intermediate level, with 53.3 % and followed by upper-intermediate level, with 20%, lower-intermediate level, with 13.3 % and advanced and mixed proficiency level students were with another 6.7%. Language environment was mostly English as a Foreign Language (EFL) classes (n=27), while a small percentage of the studies were in English as a Second Language (ESL) environment (n=2). Some of the studies did not mention the language environment but were in a global context (n=1). Conducting so many feedback studies in the EFL context has shown that feedback is a global pedagogical concern in teaching the English language, supporting feedback-related previous studies (Abdel Latif et al., 2024; Herra & Kulińska, 2018; Hidayat et al., 2024).

When the education level of the participants in these studies is considered, a significant majority of the studies were conducted with students at the tertiary level (n=26), while a small number of studies focused on secondary school students (n=4). Taking all these findings together, the research on comparative feedback studies explored English, intermediate, and higher education contexts. However, advanced-level students or ESL contexts have not been investigated deeply. Despite important aspects for the functions of feedback in EFL and intermediate level contexts, the findings cannot be generalized for all English language classrooms. To provide a more comprehensive understanding of the functions of AI vs teacher-provided feedback, future research should expand into a more diverse educational and linguistic setting.

Table 1.

Research Contexts in Comparative Studies in Relation to Teacher vs AI Feedback

Dimension	Subcategory	No. of Studies	Percent
Proficiency level	Lower-intermediate (A2–B1)	4	13.3%
	Intermediate (B1–B2)	16	53.3%
	Upper-intermediate (B2–C1)	6	20.0%
	Advanced (C1)	2	6.7%
	Mixed proficiency	2	6.7%
Language environment	EFL (English as a Foreign Language)	27	90.0%
	ESL (English as a Second Language)	2	6.7%
	Global / Not specified	1	3.3%
Education level	Secondary (high school)	4	13.3%

Tertiary (undergraduate / higher education, including distance or open education)	26	86.7%
---	----	-------

3.2 Feedback Type

One area of interest was the focus and comprehension of feedback provided by the AI tools and the teachers in these studies. As illustrated in Table 2a, the focus of the feedback provided by AI tools and teachers in comparative studies differs according to the feedback source. AI-provided feedback mostly focused only on form (n=16) (e.g., Sanosi, 2022; Thi & Nikolov, 2023; Ajabshir & Ebadi, 2023; Yıldız & Kuru Gonen, 2024), partially on form and meaning (n=10) (e.g., Steiss et al., 2024; Alsofyani & Barzanji, 2025; Zou et al., 2025), and only a small percentage of the reported studies focused on form and meaning at the same time (n=4). On the contrary, teacher-provided feedback focused on form and meaning at the same time in almost every study (n=25). Importantly, none of the feedback driven by the teachers was focused on form only. Yet, a small percentage focused on form and partially meaning (n=5). This highlights the holistic nature of teacher input.

Table 2a.

Feedback source	Form-focused only	Form + partial meaning	Comprehensive (form + meaning)	Total
AI Feedback (n=30)	16 (53.3%)	10 (33.3%)	4 (13.3%)	30
Teacher Feedback (n=30)	0 (0%)	5 (16.7%)	25 (83.3%)	30
<i>Focus and Comprehension of the Feedback Provided by the AI Tools and the Teachers</i>				

As summarised in the table above, this analysis is consistent with previous reviews of the Automated Writing Evaluation (AWE), which has long highlighted that automated tools tend to focus on superficial corrections (Stevenson and Phakiti, 2014; Zhang and Hyland, 2018).

Additionally, when AI tools used in relation to the comparison of AI-mediated feedback to teacher feedback were identified, 11 different tools stood out (see Table 2b).

Table 2b.*AI Tools Used in Relation to the Comparison of AI-Mediated Feedback to Teacher Feedback*

AI Tool (n = 11)	Type / Focus	Studies (30 total)	No. of Studies (%)
ChatGPT (OpenAI)	LLM – comprehensive, adaptive feedback on form + meaning	Guo & Wang (2023); Yan (2024); Alsofyani & Barzanji (2025); Asadi et al. (2025); Yıldız & Kuru Gonen (2024); Lo et al. (2025); Abdi Tabari et al. (2025); Tran (2025); Kao & Reynolds (2024); Steiss et al. (2024); Alnemrat et al. (2025); Yıldız et al. (2025); Fan (2023)	13 (43.3 %)
Pigaiwang	Chinese AWE – sentence-level grammar + vocabulary feedback	Wang & Han (2022)	1 (3.3 %)
Criterion (ETS)	AWE – grammar, usage, mechanics, holistic scoring	Sari & Han (2024); Han & Sari (2022); Taskiran & Goksel (2022)	3 (10.0 %)
Grammarly	Commercial AWE – grammar, vocabulary, clarity, mechanics	Ajabshir & Ebadi (2023); Sanosi (2022); Thi & Nikolov (2023); Wang (2024)	4 (13.3 %)
Write & Improve (Cambridge English)	Online AWE – automated scoring and form feedback	Lin & Crosthwaite (2024); Zhang & Zhang (2022)	2 (6.7 %)
WriteToLearn (Pearson)	AWE – content + organization scoring and revision support	Henderson et al. (2025)	1 (3.3 %)
MyAccess (Vantage Learning)	AWE – analytic and holistic scoring for EFL writing	Escalante et al. (2023)	1 (3.3 %)
eRevise (ETS family)	AWE – formative writing practice tool	Escalante et al. (2023) [same study reported dual system use]	1 (3.3 %)
Custom LLM Peer-Feedback Platform	Research-built AI peer feedback system	Zou et al. (2025)	1 (3.3 %)
AI Metalinguistic Feedback Tool	Locally developed AI providing rule-based metalinguistic feedback	Luo et al. (2025)	1 (3.3 %)
Institutional / Hybrid AI Writing Platform	University or project-specific AI assistant integrating teacher input	Wale & Kassahun (2024); Thi et al. (2022); Zhang et al. (2025)	3 (10.0 %)

When the timeline of the selection of the tools was considered, AWE tools were used before 2023 mostly, whereas after 2023, Generative AI tools started to become popular. This trend changed the content of feedback, where hybrid teacher and AI feedback is found to simulate human-like answers.

This finding aligns with Shi and Aryadoust's (2023) stance that the writing feedback construct is developing with the integration of AI. According to the dataset in the current review, nearly half of the studies using generative AI tools reported more comprehensive coverage, while only 13% of traditional AWE studies was found to provide comprehensive feedback.

3.3 Learner Response

While response to feedback varied, the majority (63.3%) of learners indicated a balanced or mixed level of trust, supporting the pedagogical proposition that AI and teacher feedback are complementary rather than competitive. As presented in Table 3, a small number of studies (10%) found that students had significantly more trust in teachers, while a larger number (26.7%) of previous work reported a stronger trust in AI, often due to speed, accessibility, or emotional benefits (Asadi et al., 2025; Zou et al., 2025).

Table 3.

Learner Response to the Feedback Source

Category	Studies	Count (n=30)	Percent
AI trusted / preferred	Alsofyani & Barzani, 2025; Asadi et al., 2025; Guo & Wang, 2023; Kao & Reynolds, 2024; Lo et al., 2025; Sari & Han, 2024; Yan, 2024; Zou et al., 2025	8	26.7%
Teacher trusted / preferred	Taskiran & Goksel, 2022; Tran, 2025; Yıldız et al., 2025	3	10.0%
Mixed / Equal trust	Abdi Tabari et al., 2025; Ajabshir & Ebadi, 2023; Alnemrat et al., 2025; Fan, 2023; Han & Sari, 2022; Henderson et al., 2025; Kao & Reynolds, 2024; Lin & Crosthwaite, 2024; Lo et al., 2025; Luo et al., 2025; Sanosi, 2022; Steiss et al., 2024; Thi & Nikolov, 2023; Wale & Kassahun, 2024; Wang & Han, 2022; Yıldız & Kuru Gonen, 2024; Zhang & Zhang, 2022; Zhang et al., 2025; Zou et al., 2025	19	63.3%

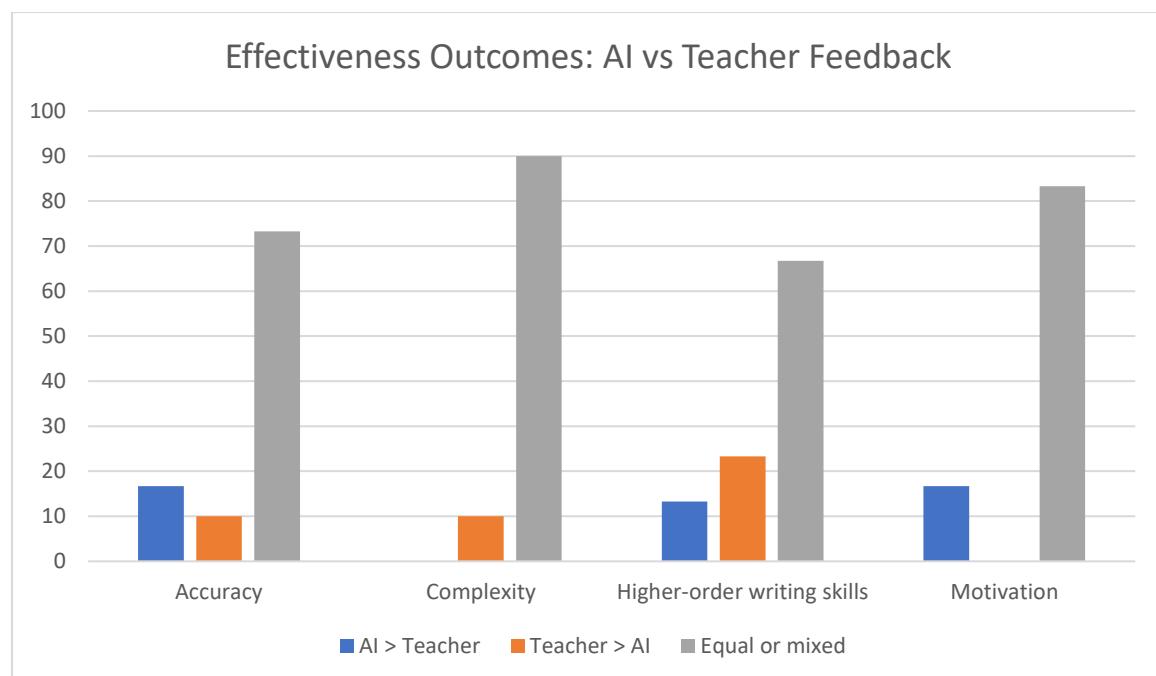
What is more, some students tended to trust AI-provided feedback more than the teachers', even if the same form of feedback was given by the teachers (Kao & Reynolds, 2024).

3.4 Effectiveness In Improving Writing

We examined the effectiveness of improving writing from four perspectives: accuracy, complexity, higher-order writing outcomes, and motivation. In all categories, equal or mixed AI and teacher feedback was found to be the most effective as illustrated in the Figure below:

Figure 2.

Effectiveness Outcomes: AI vs Teacher Feedback



Within the scope of this review, the effect of feedback on grammar, vocabulary, and mechanics, namely surface-level accuracy, was investigated. As seen in Table 4a, no significant difference between AI and teacher feedback was reported by most of the studies (73.3%). It was observed that AI tools, especially those improved for grammar checking, like Grammarly, proved superior to teachers on a small minority of the studies (16.7%) (e.g., Sari & Han, 2024; Guo & Wang, 2023), supporting previous findings of AWE (Stevenson and Phakiti, 2014; Shi and Aryadoust, 2023). However, teachers still had a role in improving writing when meaning supported correct expression (10%).

Table 4a.*Accuracy*

Category	Studies	Count (n=30)	Percent
AI	> Ajabshir & Ebadi, 2023; Alsofyani & Barzanji, 2025; Guo & Wang, 2023; Sari & Han, 2024; Wang & Han, 2022	5	16.7%
Teacher	Han & Sari, 2022; Taskiran & Goksel, 2022; Zhang & Zhang, 2022	3	10.0%
Teacher > AI			
Equal	/ Abdi Tabari et al., 2025; Alnemrat et al., 2025; Asadi et al., 2025; Fan, 2023; Henderson et al., 2025; Kao & Reynolds, 2024; Lin & Crosthwaite, 2024; Lo et al., 2025; Sanosi, 2022; Sari & Han, 2024; Steiss et al., 2024; Thi & Nikolov, 2023; Thi et al., 2022; Wale & Kassahun, 2024; Yan, 2024; Yıldız & Kuru Gonen, 2024; Yıldız et al., 2025; Zhang et al., 2025; Zou et al., 2025	22	73.3%
Mixed			

The effect of feedback on improving writing skills regarding the level of complexity was examined by looking at its effect at syntactic and lexical levels (see Table 4b). The findings showed no significant difference between teacher and AI feedback (90%). None of the studies reported that AI feedback was more effective than teacher feedback in terms of detecting complexity.

Table 4b*Complexity*

Category	Studies	Count (n=30)	Percent
AI > Teacher	–	0	0%
Teacher > AI	Han & Sari, 2022; Thi & Nikolov, 2023; Zhang & Zhang, 2022	3	10.0%
Equal	/ Abdi Tabari et al., 2025; Ajabshir & Ebadi, 2023; Alnemrat et al., 2025; Alsofyani & Barzanji, 2025; Asadi et al., 2025; Fan, 2023; Henderson et al., 2025; Kao & Reynolds, 2024; Lin & Crosthwaite, 2024; Lo et al., 2025; Luo et al., 2025; Sanosi, 2022; Sari & Han, 2024; Steiss et al., 2024; Taskiran & Goksel, 2022; Thi et al., 2022; Wale & Kassahun, 2024; Wang & Han, 2022; Yan, 2024; Yıldız & Kuru Gonen, 2024; Yıldız et al., 2025; Zhang et al., 2025; Zou et al., 2025; Tran, 2025	27	90.0%
Mixed			

The effect of feedback in improving writing related to higher-order writing outcomes was investigated concerning organization, coherence, argumentation, and genre quality. As can be seen in Table 4c, the analysed studies either reported no significant difference, or hybrid models being the most effective ones in most of the studies (66.7%). AI feedback outperformed teacher feedback in only four studies, while teacher feedback led to stronger improvements in organization, coherence, and argumentation (23.3%). Overall, teacher-provided feedback appeared more effective for higher-order improvements; however, recent findings suggest that AI feedback can complement teacher guidance when carefully integrated, which calls for a collaborative feedback model.

Table 4c.

Higher-Order Writing Outcomes

Category	Studies	Count (n=30)	Percent
AI > Teacher	Asadi et al., 2025; Guo & Wang, 2023; Steiss et al., 2024; Zou et al., 2025	4	13.3%
Teacher > AI	Ajabshir & Ebadi, 2023; Lin & Crosthwaite, 2024; Lo et al., 2025; Thi & Nikolov, 2023; Wang & Han, 2022; Yıldız et al., 2025; Zhang & Zhang, 2022	7	23.3%
Equal / Mixed (hybrid most effective or no significant difference)	Abdi Tabari et al., 2025; Alnemrat et al., 2025; Alsofyani & Barzanji, 2025; Escalante et al., 2023; Fan, 2023; Han & Sari, 2022; Henderson et al., 2025; Kao & Reynolds, 2024; Luo et al., 2025; Sanosi, 2022; Sari & Han, 2024; Taskiran & Goksel, 2022; Thi et al., 2022; Tran, 2025; Wale & Kassahun, 2024; Wang, 2024; Yan, 2024; Yıldız & Kuru Gonen, 2024; Zhang et al., 2025	19	66.7%

As illustrated in Table 4d, the effect of feedback on motivation was found to be more effective in boosting motivation and self-confidence and reducing anxiety if the feedback was a combination of two sources, AI tools and teachers (83.3%). Learner motivation boosted when AI-provided feedback was superior to teacher feedback (16.7%). This finding aligned with the previous studies searching for the effect of AI-mediated written feedback on motivation (Ahmed et al., 2025; Huang & Mizumoto, 2024; Mohammed & Khalid, 2025; Song & Song, 2023). None of the studies reported that teacher-provided feedback affected learner motivation positively.

Table 4d.*Motivation*

Category	Studies	Count	Percent (n= 30)
AI >	Alsofyani & Barzanji, 2025; Guo & Wang, 2023; Lo et al., 2025;	5	16.7%
Teacher	Sari & Han, 2024; Tran, 2025		
Teacher	-	0	0%
> AI			
Equal /	Abdi Tabari et al., 2025; Ajabshir & Ebadi, 2023; Alnemrat et al., 2025; Asadi et al., 2025; Escalante et al., 2023; Fan, 2023; Han & Sari, 2022; Henderson et al., 2025; Kao & Reynolds, 2024; Lin & Crosthwaite, 2024; Luo et al., 2025; Sanosi, 2022; Steiss et al., 2024; Taskiran & Goksel, 2022; Thi et al., 2022; Thi & Nikolov, 2023; Wale & Kassahun, 2024; Wang, 2024; Wang & Han, 2022; Yan, 2024; Yildiz & Kuru Gonen, 2024; Yildiz et al., 2025; Zhang & Zhang, 2022; Zhang et al., 2025; Zou et al., 2025	25	83.3%

3.5 Pedagogical Implications

When pedagogical implications in studies comparing AI-provided feedback to teacher-provided feedback were identified, five common implications emerged as given in Table 5 below.

Table 5.*Pedagogical Implications of Comparative Studies Regarding AI vs Teacher-Provided Feedback*

Pedagogical Implication	Studies	Count (n=30)	Percent
Hybrid feedback recommended (AI + teacher together more effective)	Abdi Tabari et al., 2025; Alnemrat et al., 2025; Alsofyani & Barzanji, 2025; Asadi et al., 2025; Escalante et al., 2023; Guo & Wang, 2023; Han & Sari, 2022; Henderson et al., 2025; Kao & Reynolds, 2024; Lin & Crosthwaite, 2024; Lo et al., 2025; Sari & Han, 2024; Steiss et al., 2024; Tran, 2025; Wang & Han, 2022; Yıldız & Kuru Gonen, 2024; Yıldız et al., 2025; Zou et al., 2025	18	60.0%
AI best for surface-level accuracy & motivational support	Ajabshir & Ebadi, 2023; Alsofyani & Barzanji, 2025; Fan, 2023; Kao & Reynolds, 2024; Lin & Crosthwaite, 2024; Lo et al., 2025; Sanosi, 2022; Sari & Han, 2024; Tran, 2025; Wang & Han, 2022	10	33.3%
Teacher feedback central for higher-order skills (organization, coherence, argumentation, genre)	Han & Sari, 2022; Henderson et al., 2025; Lin & Crosthwaite, 2024; Taskiran & Goksel, 2022; Thi & Nikolov, 2023; Yıldız et al., 2025; Zhang & Zhang, 2022; Zhang et al., 2025	8	26.7%
Learner trust & perception shape effectiveness	Alnemrat et al., 2025; Escalante et al., 2023; Guo & Wang, 2023; Henderson et al., 2025; Kao & Reynolds, 2024; Lo et al., 2025; Wale & Kassahun, 2024; Wang, 2024; Yan, 2024; Yıldız & Kuru Gonen, 2024	10	33.3%
Teacher training & institutional support needed	Abdi Tabari et al., 2025; Escalante et al., 2023; Guo & Wang, 2023; Henderson et al., 2025; Lin & Crosthwaite, 2024; Luo et al., 2025; Thi et al., 2022; Wale & Kassahun, 2024	8	26.7%

Note. Percentages are calculated with n=30; categories are non-exclusive; totals do not sum to 100%.

The reached conclusion was that combining AI and teacher feedback would yield more effective results (60%). Two other supporting conclusions drawn were that teacher feedback was considered central to higher-order skills (26.7%) while AI feedback was effective for surface-accuracy and had a positive impact on motivation (33.3%), suggesting AI-assisted feedback could be a good complement to teacher feedback. Additionally, it can be implied that learner trust and perception shape the effectiveness of feedback (16.7%). Finally, teacher training and institutional support were reported to be necessary in educational contexts (26.7%).

4. Conclusions

This systematic review synthesized findings from 30 empirical studies comparing feedback provided by AI tools and human teachers in five key dimensions: research context, feedback type, learner response, effectiveness in improving writing, and pedagogical implications for classroom practice. Since the number of papers have been skyrocketing on the subject, especially after 2023, to give an up-to-date record of recent developments, papers published in high impact journals within the last four years were included in the analysis criteria.

When the context of the studies was considered, most research was conducted in higher education EFL settings, and students were mostly intermediate level, demonstrating that AI feedback was generally situated in academic writing instruction. This limited context highlights the need for further work on the effectiveness of AI assisted feedback as opposed to human teacher feedback on lower age groups like secondary and high school students with a higher level of language proficiency. The findings of this review shed light on the scope of the master's thesis of the first author who has been investigating the role of the feedback agent (AI vs. human teachers) and the type of essay (argumentative vs. cause) on the perception and development of writing skills of highschoolers with an upper intermediate level of language proficiency. Future research needs to extend to secondary and ESL contexts and investigate the long-term effects of AI-assisted feedback on writing development.

Second, feedback type varied according to the feedback source. AI-provided feedback mostly focused on form, whereas teacher-provided feedback was more comprehensive, focusing on both form and meaning. Learners appreciated AI's immediacy, accessibility, and non-judgmental tone, which contributed to their motivation positively. On the other hand, teacher feedback offered more comprehensive support, especially for higher-order writing skills. Recent studies are increasingly drawing attention to the value of hybrid feedback models, where AI tools complement rather than replace teacher feedback to address both surface accuracy and overall writing quality.

Finally, integrating pedagogically effective AI tools necessitates teacher training, feedback literacy for both pre- and in-service teachers and students, along with institutional support for ethical use. Considering what has been found today, the role of AI feedback agents cannot be disregarded since many findings reported hybrid feedback as the most beneficial approach in second language writing development. Overall, the future of writing feedback is likely to be shaped by the collaboration of AI-teacher systems that combine technological effectiveness with human judgment and contextual understanding.

Acknowledgement

This study is based on the first author's M.A. thesis titled "Can AI give better second language written corrective feedback than teachers?: A case study from Turkish high schoolers" and supervised by the second author.

References

Abdel Latif, M. M., Alsahil, A., & Alsuhaimi, Z. (2024). Profiling EFL writing teachers' feedback provision practices and activity uses in Saudi universities. *Humanities and Social Sciences Communications*, 11(1). <https://doi.org/10.1057/s41599-024-03639-5>

Abdi Tabari, M., Kushki, A., & Wang, Y. (2025). Comparing the effects of teacher- and AI-mediated corrective feedback on accuracy, complexity, and quality in L2 written narratives. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2025.2561608>

Ahmed, I., Ghafoor, A., Liliuara, D., & Akyuningrum, V. Q. (2025). The impact of AI learning tools on ESL learners' motivation and success: A systematic literature review. *English Learning Innovation*, 6(1), 134-142. <https://doi.org/10.22219/englie.v6i1.38961>

Ajabshir, Z. F., & Ebadi, S. (2023). The effects of automatic writing evaluation and teacher-focused feedback on CALF measures and overall quality of L2 writing across different genres. *Asian-Pacific Journal of Second and Foreign Language Education*, 8(1). <https://doi.org/10.1186/s40862-023-00201-9>

Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlShareefen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1614673>

Alsofyani, A., & Barzani, S. (2025). The effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: A mixed-methods study. *Journal of Educational Computing Research*, 63(2), 215–236. <https://doi.org/10.1177/07356331241324153>

Asadi, M., Ebadi, S., & Mohammadi, L. (2025). The impact of integrating ChatGPT with teachers' feedback on EFL writing skills. *Thinking Skills and Creativity*, 56, 101766. <https://doi.org/10.1016/j.tsc.2025.101766>

Duman, G., Orhon, G. & Gedik, N. (2015) Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, 27(2): 197–216. <https://doi.org/10.1017/S0958344014000287>

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00425-2>

Fan, N. (2023). Exploring the effects of automated written corrective feedback on EFL students' writing quality: A mixed-methods study. *Sage Open*, 13(2). <https://doi.org/10.1177/21582440231181296>

Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29(7), 8435–8463. <https://doi.org/10.1007/s10639-023-12146-0>

Han, T., & Sari, E. (2022). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning*, 37(4), 961–985. <https://doi.org/10.1080/09588221.2022.2067179>

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>

Henderson, M., Bearman, M., Chung, J., Fawns, T., Buckingham Shum, S., Matthews, K. E., & de Mello Heredia, J. (2025). Comparing generative AI and teacher feedback: Student perceptions of usefulness and trustworthiness. *Assessment & Evaluation in Higher Education*. Advance online publication. <https://doi.org/10.1080/02602938.2025.2502582>

Herra, A., & Kulińska, A. (2018). The role of feedback in the process of learning English as a foreign language. *Forum Filologiczne Ateneum*, (1(6)2018), 127-143. [https://doi.org/10.36575/2353-2912/1\(6\)2018.127](https://doi.org/10.36575/2353-2912/1(6)2018.127)

Hidayat, M., Umalihayati, U., Hilman, H., & Rachman, A. (2024). Feedback challenges and strategies in EFL classroom: Insights from non-formal education institution. *English Learning Innovation*, 5(1), 86-97. <https://doi.org/10.22219/englie.v5i1.30295>

Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics*, 1, 102324. <https://doi.org/10.29140/dal.v1.102324>

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101. <https://doi.org/10.1017/S0261444806003399>

Kao, C. W., & Reynolds, B. L. (2024). Timed second language writing performance: Effects of perceived teacher vs perceived automated feedback. *Humanities and Social Sciences Communications*, 11, 1012. <https://doi.org/10.1057/s41599-024-03522-3>

Lin, S., & Crosthwaite, P. (2024). The grass is not always greener: Teacher vs. GPT-assisted written corrective feedback. *System*, 127, 103529. <https://doi.org/10.1016/j.system.2024.103529>

Lo, N., Chan, S., & Wong, A. (2025). Evaluating teacher, AI, and hybrid feedback in English language learning: Impact on student motivation, quality, and performance in Hong Kong. *Sage Open*, 15(3), 1-16. <https://doi.org/10.1177/21582440251352907>

Luo, M., Hu, X., & Zhong, C. (2025). The collaboration of AI and teacher in feedback provision and its impact on EFL learner's argumentative writing. *Education and Information Technologies*, 30(12), 17695-17715. <https://doi.org/10.1007/s10639-025-13488-7>

Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: Mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia*, 15(1). <https://doi.org/10.1186/s40468-025-00343-2>

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>

Phan, T. M., & Hoang, V. P. (2025). Enhancing EFL writing revision practices: The impact of AI- and teacher-generated feedback and their sequences. *Education Sciences*, 15(2), 232. <https://doi.org/10.3390/educsci15020232>

Sanosi, A. (2022). To err is human: Comparing human and automated corrective feedback. *Information Technologies and Learning Tools*, 90(4), 149–161. <https://doi.org/10.33407/itlt.v90i4.4980>

Sari, E., & Han, T. (2024). The impact of automated writing evaluation on English as a foreign language learners' writing self-efficacy, self-regulation, anxiety, and performance. *Journal of Computer Assisted Learning*, 40(5), 2065–2080. <https://doi.org/10.1111/jcal.13004>

Shi, H., & Aryadoust, V. (2023). *A systematic review of automated writing evaluation systems*. *Education and Information Technologies*, 28(1), 771-795. <https://doi.org/10.1007/s10639-022-11200-7>

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1260843>

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>

Stevenson, M., & Phakiti, A. (2014). *The effects of computer-generated feedback on the quality of writing*. *Assessing Writing*, 19(2), 51-65. <https://doi.org/10.1016/j.asw.2013.11.007>

Taskiran, A., & Goksel, N. (2022). Automated feedback and teacher feedback: Writing achievement in learning English as a foreign language at a distance. *Turkish Online Journal of Distance Education*, 23(2), 120–139. <https://doi.org/10.17718/tojde.1096260>

Thi, N. K., & Nikolov, M. (2023). Effects of teacher, automated, and combined feedback on syntactic complexity in EFL students' writing. *Asian-Pacific Journal of Second and Foreign Language Education*, 8(1). <https://doi.org/10.1186/s40862-022-00182-1>

Thi, N. K., Nikolov, M., & Simon, K. (2022). Higher-proficiency students' engagement with and uptake of teacher and Grammarly feedback in an EFL writing course. *Innovation in Language Learning and Teaching*, 17(3), 690-705. <https://doi.org/10.1080/17501229.2022.2122476>

Tran, T. T. (2025). Enhancing EFL writing revision practices: The impact of AI- and teacher-generated feedback and their sequences. *Education Sciences*, 15(2), 232. <https://doi.org/10.3390/educsci15020232>

Wale, B. D., & Kassahun, Y. F. (2024). The transformative power of AI writing technologies: Enhancing EFL writing instruction through the integrative use of Writerly and Google Docs. *Human Behavior and Emerging Technologies*, 2024, 9221377. <https://doi.org/10.1155/2024/9221377>

Wang, D. (2024). Teacher- Versus AI-generated (Poe application) corrective feedback and language learners' writing anxiety, complexity, fluency, and accuracy. *The International Review of Research in Open and Distributed Learning*, 25(3), 37-56. <https://doi.org/10.19173/irrodl.v25i3.7646>

Wang, Z., & Han, F. (2022). The effects of teacher feedback and automated feedback on cognitive and psychological aspects of foreign language writing: A mixed-methods research. *Frontiers in Psychology*, 13, 909802. <https://doi.org/10.3389/fpsyg.2022.909802>

Yan, D. (2024). Comparing individual vs. collaborative processing of chatgpt-generated feedback: Effects on L2 writing task improvement and learning. *Language Learning & Technology*, 28(1), 1-19. <https://doi.org/10.64152/10125/73597>

Yıldız, H., & Kuru Gonen, S. I. (2024). Automated writing evaluation system for feedback in the digital world: An online learning opportunity for English as a foreign language students. *Turkish Online Journal of Distance Education*, 25(3), 183–206. <https://doi.org/10.17718/tojde.1169727>

Yıldız, M., Topuz, A. C., Polat, H., Taşlıbeyaz, E., Kurşun, E., & Yeşilyurt, S. (2025). A comparative analysis of human and GenAI-generated feedback on EFL students' argumentative writing performance. *Educational Psychology. Advance online publication*. <https://doi.org/10.1080/01443410.2025.2539797>

Zhang, J., & Zhang, L. J. (2022). The effect of feedback on metacognitive strategy use in EFL writing. *Computer Assisted Language Learning*, 37(5–6), 1198–1223. <https://doi.org/10.1080/09588221.2022.2069822>

Zhang, Z. V., & Hyland, K. (2018). *Student engagement with teacher and automated feedback on L2 writing*. Assessing Writing, 36, 90-102. <https://doi.org/10.1016/j.asw.2018.02.004>

Zhang, Z., Aubrey, S., Huang, X., & Chiu, T. K. (2025). The role of generative AI and hybrid feedback in improving L2 writing skills: A comparative study. *Innovation in Language Learning and Teaching*, 1-19. <https://doi.org/10.1080/17501229.2025.2503890>

Zou, S., Guo, K., Wang, J., & Liu, Y. (2025). Investigating students' uptake of teacher- and ChatGPT-generated feedback in EFL writing: A comparison study. *Computer Assisted Language Learning. Advance online publication*. <https://doi.org/10.1080/09588221.2024.2447279>

AUTHOR BIODATA

Aysun Barış Kosaogulları completed her B.A. in Teaching English as a Foreign Language at Middle East Technical University. She is currently pursuing her M.A. in English Language Teaching at İzmir Democracy University and works as an English teacher at İzmir Girls High School.

Vasfiye Geçkin completed her B.A. and M.A. in Teaching English as a Foreign Language at Boğaziçi University, Türkiye and her Ph.D degree in Linguistics from Universität Potsdam (Germany) and Macquarie University (Australia). Currently, she works as an Assistant professor at the Department of Foreign Language Education in Izmir Democracy University, Türkiye.

Teaching Prosody Through Integrated Skills: An AI-Supported Pedagogical Perspective

Prof. Dr. Güven Mengü⁷, Ankara Haci Bayram Veli University,
Department of English Language and Literature, Ankara, Türkiye

APA Citation:

Mengü G. (2025). Teaching Prosody Through Integrated Skills: An AI-Supported Pedagogical Perspective

Abstract

Prosody plays a central role in spoken communication, shaping how listeners understand, process, and interpret speech. Even when grammatical accuracy and vocabulary range are relatively strong, difficulties with rhythm, stress, and intonation can make speech hard to follow or pragmatically inappropriate. Listener-based research has repeatedly shown that suprasegmental features contribute more directly to intelligibility than many segmental deviations, particularly in real-time interaction (Derwing & Munro, 2009). Despite this evidence, prosody continues to occupy a marginal position in many English Language Teaching contexts, where it is often treated as a minor pronunciation concern or left to develop implicitly. This paper argues that recent advances in artificial intelligence create new pedagogical conditions for addressing this long-standing gap. Rather than proposing technology as a stand-alone solution, the discussion situates AI-supported prosody instruction within an integrated skills framework, where listening, speaking, reading, and writing are treated as interconnected processes. Drawing on research in intelligibility-oriented pronunciation teaching, cognitive theories of noticing and output, and sociocultural perspectives on mediated learning, the paper explores how tools such as Automatic Speech Recognition, neural Text-to-Speech, and AI-based prosody visualization can support sustained prosodic development. It is argued that when these tools are guided by clear pedagogical goals and interpreted through teacher mediation, they can make prosodic features more visible, more practicable, and more meaningful for learners. The paper concludes that AI-supported integration offers a realistic and scalable way of repositioning prosody as a core component of communicative competence rather than a peripheral skill.

Keywords: prosody, suprasegmentals, intelligibility, integrated skills, artificial intelligence, automatic speech recognition (ASR), neural text-to-speech (TTS), prosody visualization, pronunciation pedagogy, communicative competence

1. Introduction

In spoken interaction, listeners rarely evaluate language ability by focusing solely on grammatical accuracy. What often matters more is how easily speech can be followed and how naturally it fits the communicative situation. These judgments are strongly influenced by prosody, which organizes speech through rhythm, stress, intonation, and pausing. When prosodic patterns align with listener expectations, speech tends to feel fluent and coherent; when they do not, even structurally accurate speech can sound effortful or confusing. Research on intelligibility has consistently shown that suprasegmental features play a decisive role in listener understanding. Early work by Anderson-Hsieh, Johnson, and Venkatagiri (1992) demonstrated that deviations in prosody often led to lower comprehensibility ratings than segmental errors. Later studies refined this observation by showing that prosodic irregularities increase listener effort, requiring more cognitive resources to process speech in

⁷ Prof. Dr. Güven Mengü
guven.mengu@hbv.edu.tr

real time (Derwing & Munro, 2009). From the listener's perspective, this effort is not always registered as outright misunderstanding but as fatigue, uncertainty, or a sense that communication is less smooth than it could be.

Despite these findings, prosody has struggled to find a stable place in English Language Teaching. In many instructional contexts, pronunciation teaching focuses on individual sounds, while rhythm and intonation receive limited attention. As Levis (2005) notes, segmental features are often easier to explain, practice, and assess, which makes them more attractive targets in classrooms constrained by time and curricular demands. Prosody, by contrast, is frequently described in general terms without clear instructional pathways, leading teachers to address it sporadically or indirectly. The challenge is particularly pronounced for learners whose first languages differ from English in rhythmic organization. English relies heavily on stress timing and vowel reduction, whereas many other languages distribute prominence more evenly across syllables. Learners from such backgrounds may produce speech that is grammatically accurate but rhythmically flat, a pattern that has been shown to increase processing load for listeners (Munro & Derwing, 1995). Without explicit support, these learners may not realize that rhythm and stress are contributing to communication difficulties.

Another factor contributing to the neglect of prosody is its perceptual complexity. Prosodic features unfold over time and are closely tied to discourse context, making them difficult to isolate and describe. Schmidt's (1990) Noticing Hypothesis helps explain why exposure alone is often insufficient: if learners do not consciously notice prosodic contrasts, they are unlikely to incorporate them into their developing language system. In classroom settings, teachers may model intonation or encourage repetition, but learners often lack the perceptual tools needed to understand what they are expected to change. Practical constraints further complicate the situation. Teaching prosody effectively requires repeated practice and individualized feedback, conditions that are difficult to meet in large classes. Many teachers, particularly those working in foreign-language contexts, also report limited confidence in evaluating suprasegmental features, which can lead to reliance on mechanical drills or to avoidance of the topic altogether (Murphy, 2014). Over time, this combination of cognitive, pedagogical, and institutional factors allows prosodic difficulties to persist and, in some cases, fossilize (Han, 2004).

Recent developments in Artificial Intelligence invite a reconsideration of these limitations. Modern AI-based tools can analyze spoken language in real time, provide immediate feedback, and visually represent prosodic features, previously accessible only through specialist phonetic software. Research on ASR-based pronunciation practice suggests that such tools can draw learners' attention to timing, stress, and pausing through interaction rather than explicit explanation, encouraging repeated adjustment and experimentation (Liakin et al., 2017). The central argument of this paper is that these technological affordances are most effective when they are embedded within an integrated skills framework. Prosody does not belong exclusively to speaking; it shapes listening comprehension, supports reading fluency, and influences written expression through internal speech. Treating prosody as a shared resource across skills offers a more realistic reflection of language use and a more coherent basis for instruction (Oxford, 2001; Hinkel, 2006). The sections that follow develop this argument in detail. First, the paper examines why prosody has remained marginalized in ELT despite its recognized importance. It then outlines the theoretical foundations that support explicit and integrated prosody instruction. Building on this foundation, the paper explores how AI-supported tools can be used to integrate prosody across listening, speaking, reading, and writing, and then addresses assessment practices and ethical considerations. The discussion concludes by arguing that AI-

supported integration offers a practical way of restoring prosody to its rightful place in communicative language teaching.

2. Why Prosody Remains Marginal in ELT Practice

The marginal position of prosody in English Language Teaching is often explained as a practical issue, usually linked to time constraints or large class sizes. While these factors certainly play a role, they do not fully account for the persistence of the problem. Prosody has remained peripheral not because teachers doubt its importance, but because it sits uncomfortably within existing curricular structures, assessment traditions, and beliefs about what pronunciation teaching should look like. One contributing factor is the way pronunciation itself is conceptualized in many teaching contexts. Pronunciation is frequently treated as a set of discrete items that can be taught, practiced, and corrected individually. This view fits segmental features reasonably well, but it does not translate easily to suprasegmentals. Rhythm and intonation are not isolated units; they operate across phrases and entire stretches of discourse. As Celce-Murcia, Brinton, and Goodwin (2010) point out, teachers often lack clear procedures for addressing these broader patterns, even when they recognize their communicative importance.

Curricular documents tend to reinforce this uncertainty. Syllabi usually specify grammar points and lexical targets with considerable precision, while prosody is mentioned, if at all, in general terms such as natural pronunciation or appropriate intonation. Because these descriptions are vague, they offer little guidance for lesson planning. Over time, what begins as a lack of clarity becomes a lack of emphasis. Teachers focus on areas that are clearly defined and institutionally visible, while prosody remains implicit and optional. Assessment practices further deepen this imbalance. In many classrooms, pronunciation assessment relies on overall impression rather than explicit criteria. Learners may be told that their speech sounds flat or unnatural, but such comments rarely translate into concrete action. Educators often note that even experienced teachers struggle to articulate what exactly is problematic in a learner's prosody, partly because suprasegmental features resist simple right–wrong judgments. When assessment is impressionistic, instruction tends to follow the same pattern.

Methodological history also matters. Early pronunciation teaching approaches emphasized accuracy at the sound level, and although prosody was never entirely absent, it was rarely foregrounded. Later communicative approaches shifted attention toward meaning and interaction, often assuming that pronunciation would improve naturally through use. As Isaacs (2009) observes, this shift left many teachers without a clear rationale or toolkit for explicit prosody instruction. Prosody became something to be picked up rather than taught. From the learner's perspective, prosody presents additional challenges. Suprasegmental patterns are difficult to notice because they are not tied to individual words and because they unfold over time. Learners may hear differences in intonation or rhythm but cannot identify what those differences consist of. Schmidt's (1990) argument that noticing is a prerequisite for acquisition is particularly relevant here. If learners do not clearly perceive prosodic contrasts, repeated exposure alone is unlikely to lead to change. This perceptual difficulty is compounded when learners' first languages differ substantially from English in rhythmic organization. English stress timing, vowel reduction, and variable syllable length may conflict with patterns learners have internalized from their first language. Research on intelligibility suggests that such mismatches increase listener effort even when speech is grammatically accurate, a finding reported consistently in work by Munro and Derwing (1995). Yet because communication often remains possible, learners may not receive enough feedback to recognize prosody as the source of difficulty.

3. Prosody, Cognition, and AI-Mediated Learning

Prosody occupies a central position in spoken language because it organizes speech in ways that listeners rely on to interpret meaning. Rhythm, stress, intonation, and pausing shape how utterances are segmented, which elements are highlighted, and how speaker intention is inferred. From an intelligibility-oriented perspective, these suprasegmental features often matter more than the accurate production of individual sounds, especially in extended discourse. Listener-based research has repeatedly shown that prosodic irregularities increase processing effort and reduce communicative efficiency, even when speech remains grammatically accurate (Derwing & Munro, 2009). Early empirical work already pointed to the disproportionate impact of prosody on listener judgments. Anderson-Hsieh, Johnson, and Venkatagiri (1992) demonstrated that deviations in stress and intonation were stronger predictors of negative comprehensibility ratings than many segmental errors. These findings suggest that prosody functions as an organizing framework for speech perception rather than as an optional stylistic layer. When this framework is disrupted, listeners must allocate additional cognitive resources to decode meaning, which can lead to fatigue, misunderstanding, or pragmatic misinterpretation.

The role of sentence stress provides a clear illustration of this process. Stress placement guides listeners in identifying which information is new, contrastive, or backgrounded within an utterance. Experimental studies by Hahn (2004) showed that appropriate primary stress significantly improved listener recall, whereas misplaced stress interfered with comprehension, even when lexical and grammatical information was intact. Such findings highlight that prosody shapes meaning at the discourse level, not merely at the phonetic level. Cognitive theories of second language acquisition help explain why prosody is particularly resistant to change. Schmidt's (1990) Noticing Hypothesis argues that linguistic input does not become intake unless learners consciously notice relevant features. Prosody poses a challenge in this respect because it is abstract, transient, and distributed across time. Learners may perceive that speech sounds different without being able to identify what exactly distinguishes one prosodic pattern from another. This difficulty is further compounded by first-language phonological systems, which shape how rhythm and intonation are perceived.

Research on multimodal input suggests that combining auditory information with visual representations can support noticing of suprasegmental features. When learners can see pitch movement or timing patterns while listening, abstract features become more concrete. Hardison (2004) found that visual feedback facilitated both perception and production of intonation patterns, indicating that noticing can be enhanced when learners engage multiple sensory channels. This finding is particularly relevant for instructional contexts in which prosody has traditionally been addressed solely through imitation. Production-based accounts of language learning provide a complementary rationale for explicit prosody instruction. Swain's (1985) Output Hypothesis emphasizes that learners refine their linguistic system when they are pushed to produce language that is precise and intelligible. For prosody, this means creating conditions in which learners must adjust their rhythm, stress, or intonation to be understood. Passive exposure to correct models is unlikely to be sufficient if learners are not required to test and revise their own production.

Automatic Speech Recognition creates such conditions by introducing a communicative constraint. When learners interact with ASR systems, recognition success depends on intelligibility rather than on grammatical correctness alone. Recognition breakdowns often draw attention to timing, pausing, or prominence rather than to individual sounds. Studies of ASR-based pronunciation practice report that learners repeatedly attempt to modify their speech, gradually adjusting suprasegmental

features in response to recognition outcomes (Liakin et al., 2017). This process aligns closely with the notion of pushed output, as learners are compelled to move beyond meaning-focused production toward form-focused refinement. Sociocultural theory offers additional insight into how technology can support prosodic development. From this perspective, learning is mediated by tools that shape activity and cognition. AI-based pronunciation tools function as mediational artifacts that support learners as they operate within their Zone of Proximal Development. By providing immediate, individualized feedback, these tools allow learners to attempt forms that would otherwise remain out of reach. McCrocklin (2016) argues that such tools can promote learner autonomy by enabling sustained, self-directed practice, particularly in areas like pronunciation where repetition is essential.

At the same time, sociocultural approaches emphasize that mediation is never neutral. Prosody carries social and pragmatic meaning that cannot be reduced to acoustic patterns alone. While AI systems can identify regularities in pitch and timing, they cannot fully interpret how these patterns function in interaction. Teachers, therefore, play a crucial role in contextualizing feedback, explaining acceptable variation, and linking prosodic choices to communicative intent. This interpretive role becomes especially important in global contexts where multiple varieties of English coexist. Taken together, these theoretical perspectives converge on a shared conclusion. Prosody is fundamental to intelligible and effective communication, yet it is difficult to acquire through exposure alone. Cognitive, output-based, and sociocultural accounts all point to the need for explicit attention, guided practice, and meaningful feedback. When AI technologies are used as mediational supports rather than as authoritative judges, they create conditions in which prosodic development becomes more visible, manageable, and sustainable for learners. Affective factors also play a role. Experimenting with new prosodic patterns can feel awkward, especially in front of peers. Unlike vocabulary errors, which are often treated as normal parts of learning, prosodic deviations may feel more personal, as they affect voice quality and expressive range. Krashen's (1982) notion of the affective filter helps explain why learners may avoid such experimentation. When anxiety is high, learners tend to retreat to familiar speech patterns, reinforcing existing habits.

Over time, these habits can become fossilized. Acton (1984) and later Han (2004) describe fossilization as a process in which certain features stabilize despite continued exposure and practice. Prosody is particularly vulnerable in this respect because it is rarely corrected in everyday communication. As long as messages get across, listeners may not signal difficulty, and learners receive little motivation to adjust their rhythm or intonation. Seen together, these factors suggest that the marginalization of prosody is not accidental. It is the result of overlapping structural, cognitive, and affective factors that make suprasegmentals difficult to teach, assess, and even discuss. Any attempt to improve prosody instruction must therefore address these underlying conditions rather than treating prosody as an add-on to existing practice.

4. Integrating Prosody Across Language Skills

Prosody does not belong to a single language skill. It operates wherever language is processed, whether learners are interpreting spoken input, producing speech, reading texts, or shaping written discourse. Treating prosody as a speaking-only concern narrows its pedagogical potential and obscures its role in comprehension and meaning-making. An integrated skills perspective enables addressing prosody as a shared resource that supports multiple dimensions of language use. In real communicative situations, skills are rarely used in isolation. Listening and speaking are closely

intertwined, just as reading and writing often involve internal speech. Oxford (2001) emphasizes that skill integration reflects how language is actually used outside the classroom, and Hinkel (2006) argues that instruction should mirror this interdependence rather than reinforce artificial boundaries. Prosody fits naturally within this view because it shapes how input is perceived and how output is organized.

Listening provides a clear example of how prosody supports comprehension. Beyond recognizing words, listeners rely on rhythm and intonation to segment speech into meaningful units. Stress placement signals informational focus, while intonation contours help listeners infer speaker attitude or intent. Research on listening comprehension has shown that learners who struggle with prosodic cues often rely heavily on lexical guessing strategies, which can compensate only partially for weak bottom-up processing (Wilson, 2003). When prosodic patterns are unfamiliar, speech may be perceived as fast or unclear even when the vocabulary is known. AI-supported listening activities help learners attend to these cues without overloading them. Visual representations of pitch movement and pausing can accompany audio input, allowing learners to notice patterns that might otherwise pass unnoticed. Levis and Pickering (2004) found that when learners were guided in interpreting intonation contours, their sensitivity to discourse-level meaning improved. Such activities shift listening from passive exposure toward active perception, encouraging learners to attend to how meaning is carried by sound as well as by words.

Reading is often assumed to be visually driven, yet prosody also plays an important role here as well. Fluent reading involves grouping words into phrases and clauses, a process closely linked to rhythm and stress. Disfluent reading, particularly in a second language, frequently reflects a word-by-word decoding strategy that ignores larger prosodic units. Rasinski (2004) notes that oral reading fluency is strongly associated with comprehension, suggesting that prosodic phrasing supports meaning construction during reading. AI-based reading tools offer new ways to support this connection. When learners read aloud to systems that track timing and pausing, they receive immediate feedback on unnatural breaks or misplaced stress. Such feedback helps learners align visual text with spoken rhythm. Over time, this alignment can also influence silent reading, as learners begin to internalize more natural phrasing patterns. Grabe (2009) describes this process as a shift from local decoding to global processing, a shift that prosody helps facilitate.

Speaking, unsurprisingly, is the area where prosody is most visible, but it is also where anxiety can most strongly inhibit experimentation. Producing unfamiliar intonation patterns requires learners to alter habitual voice use, which can feel uncomfortable in social settings. Krashen's (1982) concept of the affective filter helps explain why learners may avoid prosodic variation even when they understand its importance. High anxiety encourages safe, familiar patterns, reinforcing monotony or first-language rhythm. AI-supported speaking practice can lower this barrier by creating low-stakes environments for experimentation. Voice-based interaction with AI systems allows learners to repeat utterances, adjust delivery, and receive feedback without social pressure. Studies on online guided pronunciation practice indicate that learners are more willing to engage in repeated prosodic adjustment when feedback comes from a system rather than from peers or teachers (Thomson, 2011). This willingness supports the kind of extensive practice required for automatization. Writing may seem distant from prosody, yet the connection becomes clear when considering the role of internal speech. Skilled writers often hear their sentences as they compose, using rhythm and pacing to guide structure. Flower and Hayes (1981) describe writing as a recursive process that involves constant interaction between planning, drafting, and reviewing, much of which is mediated by internalized

speech. From this perspective, punctuation functions as a written representation of prosodic cues, signaling pauses, emphasis, and boundaries.

Neural Text-to-Speech tools make this relationship explicit by turning written text into spoken output. When learners listen to their own writing read aloud, prosodic weaknesses often become immediately apparent. Run-on sentences, awkward phrasing, or unclear emphasis are easier to detect aurally than visually. Chafe (1988) argues that listening to written language reveals how punctuation and syntax interact with prosody, a realization that can lead to more coherent revision. Integrating prosody across skills also supports consistency in learning. Rather than encountering rhythm and intonation sporadically, learners experience them repeatedly in different modalities. This repetition across contexts strengthens mental representations and reduces the likelihood that prosody will remain compartmentalized or superficial. Gilbert's (2008) prosody pyramid captures this idea by positioning rhythm as the foundation that supports grammar and vocabulary, rather than as an optional overlay. An integrated approach does not imply that every activity must address all skills simultaneously. Instead, it involves maintaining a shared focus that carries across tasks and modalities. When learners listen for stress patterns, read with attention to phrasing, speak with awareness of prominence, and revise writing by listening for rhythm, prosody becomes part of how language is processed rather than an isolated object of instruction. This integration is particularly important in contexts where exposure to natural spoken English is limited. In such settings, AI-supported tools can extend the classroom by providing consistent models and feedback across skills. The challenge lies not in the availability of technology but in aligning its use with pedagogical goals that prioritize intelligibility, meaning, and learner agency.

5. Feedback, Assessment, and the Pedagogical Role of AI

One of the persistent difficulties in teaching prosody lies in how it is evaluated and discussed with learners. Unlike grammar or vocabulary, prosodic features do not lend themselves easily to discrete scoring. Teachers often rely on holistic impressions, describing speech as flat, unnatural, or unclear, yet such labels rarely provide learners with enough information to guide improvement. As a result, feedback on prosody tends to remain vague, sporadic, and weakly connected to assessment practices. AI-supported tools introduce new possibilities by making prosodic features more observable and traceable over time. Through acoustic analysis, these systems can identify patterns in pitch range, timing, and stress distribution, providing concrete reference points for feedback. The value of this information does not lie in the numerical scores themselves, but in how patterns can be visualized and compared across attempts. Recent research emphasizes that pattern-based feedback aligns more closely with how prosody functions in communication, since meaning emerges from repeated prosodic choices rather than isolated errors.

From a formative assessment perspective, this shift is particularly important. When learners receive feedback that highlights tendencies rather than single mistakes, they are better positioned to reflect on their own speech habits. For example, consistently placing prominence on function words or maintaining a narrow pitch range becomes visible as a recurring pattern rather than an abstract comment. Research on online guided pronunciation practice suggests that such feedback encourages learners to engage in self-monitoring and revision, especially when opportunities for repetition are built into the task design (Thomson, 2011). Self-assessment also gains a more concrete basis in AI-supported environments. When learners can compare their own productions across time, they develop a clearer sense of progress and remaining challenges. This longitudinal perspective is difficult to achieve in traditional classrooms, where feedback is often ephemeral. Mora-Plaza and Rakowicz

(2023) note that sustained interaction with ASR-based systems can help learners internalize prosodic targets through repeated cycles of production and adjustment, particularly at lower proficiency levels.

At the same time, the use of AI in assessment raises important pedagogical questions. Acoustic similarity to a reference model does not automatically guarantee communicative appropriateness. Prosodic variation is shaped by context, genre, and speaker identity, and what counts as effective intonation in one situation may sound inappropriate in another. For this reason, AI-generated feedback must be interpreted rather than accepted at face value. Levis (2020) cautions that pronunciation assessment should prioritize intelligibility and listener effort over strict adherence to idealized norms. This interpretive work highlights the continuing importance of the teacher's role. Within the TPACK framework, teachers are not replaced by technology; instead, their expertise shifts toward analysis, mediation, and decision-making (Mishra & Koehler, 2006). AI tools can generate data, but teachers determine which features matter for a given communicative goal and how feedback should be framed. For instance, a rising terminal contour may be flagged as non-standard by an AI model trained on broadcast speech, yet a teacher can explain its pragmatic function in conversational English, drawing on sociolinguistic research such as Warren's (2016) work on uptalk.

Teacher mediation is also essential for maintaining learner confidence. Automated feedback, if presented without context, can appear absolute or discouraging. When teachers guide learners in interpreting results, feedback becomes a resource rather than a judgment. Murphy (2014) argues that intelligible, non-native models should be treated as legitimate targets in pronunciation teaching, a principle that applies equally to AI-supported environments. Learners benefit when feedback is framed in relation to communicative effectiveness rather than native-like perfection. Assessment practices supported by AI are therefore most effective when they remain formative and dialogic. Instead of replacing teacher evaluation, AI tools extend it by offering additional perspectives on learner performance. This extension allows classroom time to be used more strategically, shifting attention away from constant correction toward guided reflection and goal setting. Ultimately, the pedagogical value of AI-supported feedback lies in how it reshapes the relationship between practice and assessment. When learners receive immediate, interpretable information about their prosodic choices, assessment becomes part of learning rather than an endpoint. This integration supports sustained development and reduces the likelihood that prosodic difficulties will remain unnoticed or unaddressed over time.

6. Ethical Considerations, Bias, and Pedagogical Responsibility

The growing presence of AI in pronunciation teaching raises ethical questions that cannot be treated as secondary concerns. While AI-supported tools offer new possibilities for addressing long-standing challenges in prosody instruction, they also shape learners' perceptions of what counts as good or acceptable speech. These perceptions, in turn, influence learner identity, confidence, and long-term goals. One of the most frequently discussed issues is algorithmic bias. AI systems are trained on large datasets, and the characteristics of these datasets inevitably shape system behavior. In pronunciation and prosody tools, training data often privileges standardized varieties of English, such as General American or Received Pronunciation. As a result, prosodic patterns associated with other legitimate varieties may be marked as deviant or suboptimal. O'Neil (2016) argues that such biases are rarely neutral, as they reflect broader social hierarchies embedded in data selection and system design.

From an English as an International Language perspective, this raises important concerns. Jenkins (2000) emphasizes that global communication does not require convergence toward a single native-speaker norm, particularly at the suprasegmental level where variation is widespread and functional. If AI tools implicitly promote a narrow model of English prosody, they risk undermining the goal of international intelligibility. Learners may be encouraged to suppress features that are communicatively effective but do not align with the system's expectations. This issue becomes especially salient in assessment contexts. Automated feedback can appear authoritative, even when learners and teachers are aware that it is based on probabilistic models rather than absolute standards. Without careful mediation, learners may interpret AI judgments as final or objective, leading to unnecessary anxiety or unrealistic self-evaluation. Levis (2020) cautions that pronunciation assessment should always be grounded in communicative outcomes, not acoustic resemblance alone, a principle that applies directly to AI-supported feedback.

Ethical use of AI therefore depends heavily on pedagogical framing. Teachers play a crucial role in explaining what AI feedback represents and what it does not. Rather than presenting system output as a verdict, instructors can position it as one source of information that must be interpreted in relation to context, purpose, and audience. When learners understand that variation is expected and often meaningful, feedback becomes a tool for reflection rather than a measure of deficiency. Data privacy and transparency also deserve attention. Many AI-based tools rely on cloud-based processing, which involves storing or analyzing learner speech beyond the classroom. While this issue is often addressed at the institutional level, teachers should remain aware of how learner data is handled and communicate this information clearly. Trust in the learning environment depends not only on pedagogical effectiveness but also on ethical transparency. Another ethical dimension concerns accessibility and equity. AI-supported pronunciation tools can extend learning opportunities beyond classroom time, but access to such tools is uneven. Learners with limited technological resources may benefit less from AI-enhanced instruction, potentially widening existing gaps. Godwin-Jones (2021) notes that technology-enhanced language learning must be evaluated not only in terms of innovation but also in terms of who can participate and under what conditions.

Pedagogical responsibility, then, involves more than selecting effective tools. It requires ongoing reflection on how technology shapes norms, expectations, and learner identities. Teachers must balance the affordances of AI with an awareness of its limitations, ensuring that technology serves communicative goals rather than redefining them. When used thoughtfully, AI-supported prosody instruction can promote inclusivity rather than conformity. By emphasizing intelligibility, flexibility, and communicative intent, educators can help learners use AI feedback as a guide rather than a constraint. This approach aligns technological innovation with human judgment, preserving the central role of pedagogy in language education.

7. Discussion

The discussion presented in this study moves beyond the question of whether prosody matters in English language learning and instead focuses on why it continues to resist systematic instruction despite decades of research demonstrating its importance. One explanation that emerges is that prosody challenges many of the assumptions underlying traditional classroom practices. Unlike grammatical structures or lexical items, suprasegmental features do not lend themselves easily to linear presentation, discrete correction, or short-term mastery. Their development requires sustained exposure, repeated experimentation, and feedback that extends across time and communicative contexts. Within this pedagogical reality, the role of Artificial Intelligence becomes significant not

because it introduces entirely new learning principles, but because it alters the conditions under which existing principles can be enacted. Cognitive accounts of second language acquisition have long emphasized the importance of noticing, practice, and feedback, yet classroom constraints often limit how frequently these conditions can be met. AI-supported tools expand these conditions by enabling learners to engage with prosodic features repeatedly and autonomously, reducing reliance on scarce classroom time. This shift helps explain why AI-supported prosody practice has shown promise even when instructional designs remain relatively simple.

Another key issue illuminated in this discussion concerns the relationship between prosody and skill integration. Prosodic patterns do not operate independently within speaking alone; they shape how speech is segmented during listening, how text is phrased during reading, and how coherence is achieved during writing through internal speech. Treating prosody as a cross-skill resource aligns instruction more closely with real language use, where skills interact continuously rather than sequentially. From this perspective, the pedagogical value of AI lies partly in its flexibility: the same technological affordances can support perception, production, and revision across different modalities. At the same time, the discussion highlights important limits to what AI can and cannot provide. Acoustic analysis and automated feedback can reveal patterns, but they cannot interpret meaning, intention, or social appropriateness. Prosody carries pragmatic and interpersonal significance that depends on context, genre, and speaker identity. This limitation reinforces the centrality of teacher mediation. Rather than acting as the primary model of pronunciation, teachers assume the role of interpretive guides who help learners understand which prosodic features matter in particular communicative situations and why variation is often both natural and functional.

Affective and identity-related factors further complicate prosody instruction and are especially relevant in AI-mediated environments. Altering rhythm or intonation can feel more personally exposing than correcting vocabulary or grammar, and learners may resist experimentation even when they understand its importance. Low-stakes AI interaction can reduce social pressure, but only when feedback is framed as supportive rather than evaluative. If automated judgments are perceived as absolute standards, they risk reinforcing anxiety or unrealistic expectations of native-like performance. Ethical considerations, therefore, remain integral to any discussion of AI-supported prosody instruction. Algorithmic bias, access to technology, and data transparency shape who benefits from these tools and how learner speech is evaluated. Without careful pedagogical framing, AI systems trained on narrow speech models may inadvertently promote conformity rather than intelligibility. The discussion underscores that responsible integration requires teachers to actively contextualize feedback, emphasize communicative effectiveness, and resist the reification of any single prosodic norm. Taken together, these points suggest that AI-supported prosody instruction is neither a technical fix nor a methodological revolution. Its value lies in enabling existing pedagogical principles to operate more consistently, particularly in areas where traditional instruction has struggled to provide sufficient practice and feedback. The implications of this shift become clearer when the broader contribution of such an approach is considered.

8. Conclusion

This paper has argued for a reorientation of prosody from a marginal pronunciation concern to a central component of communicative competence in English language education. While the importance of suprasegmental features for intelligibility and pragmatic interpretation has been well documented, instructional practices have not consistently reflected this understanding. The analysis presented here suggests that recent developments in Artificial Intelligence offer a realistic opportunity

to address this imbalance, provided that technology is embedded within sound pedagogical principles. By supporting repeated practice, multimodal feedback, and learner autonomy, AI-supported tools make prosodic features more accessible as objects of learning. When prosody is addressed across listening, speaking, reading, and writing, learners encounter rhythm and intonation not as isolated techniques but as integral aspects of language processing and expression. This integrated perspective helps bridge the gap between classroom activities and real-world communication, where prosody continuously shapes meaning.

It is important to stress that technological support does not diminish the role of the teacher. On the contrary, effective use of AI requires informed mediation, interpretive guidance, and ethical awareness. Teachers remain responsible for aligning feedback with communicative goals, recognizing legitimate variation, and fostering learner confidence. AI functions best not as an authority on correctness but as a resource that supports reflection and sustained engagement. The contribution of this paper is therefore conceptual and pedagogical rather than empirical. It brings together research on intelligibility, cognition, skill integration, and technology-enhanced learning to propose a coherent framework for AI-supported prosody instruction. Future research is needed to examine how such approaches operate over time, how gains transfer to spontaneous interaction, and how AI systems can be developed to better reflect the diversity of global English use.

References

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly*, 18(1), 71–85. <https://doi.org/10.2307/3585752>

Anderson-Hsieh, J., Johnson, R., & Venkatagiri, H. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42(4), 529–555. <https://doi.org/10.1111/j.1467-1770.1992.tb01043.x>

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.

Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication*, 5(4), 395–426. <https://doi.org/10.1177/074108838005004001>

Derwing, T. M., & Munro, M. J. (2009). Putting a focus on pronunciation. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 379–397). Routledge.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>

Gilbert, J. B. (2008). *Teaching pronunciation: Using the prosody pyramid*. Cambridge University Press.

Godwin-Jones, R. (2021). Big data and language learning: Opportunities and challenges. *Language Learning & Technology*, 25(1), 4–19.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201–223. <https://doi.org/10.2307/3588378>

Han, Z. H. (2004). *Fossilization in adult second language acquisition*. Multilingual Matters.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training to segmental accuracy in non-native speech. *Language Learning*, 54(4), 579–620. <https://doi.org/10.1111/j.0023-8333.2004.00283.x>

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109–131. <https://doi.org/10.2307/40264513>

Isaacs, T. (2009). Integrating pronunciation work into the communicative language teaching classroom. *TESL Canada Journal*, 27(1), 1–12.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369–377. <https://doi.org/10.2307/3588485>

Levis, J. M. (2020). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

Levis, J. M., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32(4), 505–524. <https://doi.org/10.1016/j.system.2004.09.003>

Liakin, D., Cardoso, W., & Liakina, N. (2017). Learning L2 pronunciation with a mobile app: The role of automatic speech recognition. *ReCALL*, 29(2), 1–20. <https://doi.org/10.1017/S0958344016000180>

McCrocklin, S. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System*, 57, 25–42. <https://doi.org/10.1016/j.system.2015.12.002>

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.

Mora-Plaza, M., & Rakowicz, A. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation instruction: A meta-analysis. *ReCALL*, 35(2), 147–166. <https://doi.org/10.1017/S0958344022000206>

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>

Murphy, J. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System*, 42, 258–269. <https://doi.org/10.1016/j.system.2013.12.007>

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Oxford, R. L. (2001). Integrated skills in the ESL/EFL classroom. *ESL Magazine*, 6(1), 5–12.

Rasinski, T. V. (2004). *Assessing reading fluency*. Pacific Resources for Education and Learning.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

Thomson, R. I. (2011). Computer-assisted pronunciation training: Targeting the suprasegmentals. *CALICO Journal*, 28(3), 744–765. <https://doi.org/10.11139/cj.28.3.744-765>

Warren, P. (2016). *Uptalk: The phenomenon of rising intonation*. Cambridge University Press.

Wilson, M. (2003). Discovery listening. *ELT Journal*, 57(4), 335–343. <https://doi.org/10.1093/elt/57.4.335>

Integrating AI into Undergraduate Research: Exploring Students' Perceptions, Practices, and Ethics

Merve Elbirlik Tülek

Özyegin University, İstanbul, Türkiye

APA Citation:

Elbirlik Tülek, M. (2025). Integrating AI into Undergraduate Research: Exploring Students' Perceptions, Practices, and Ethics.

Abstract

This study explores undergraduate engineering students' perceptions, practices, and ethical awareness related to the use of artificial intelligence (AI) tools in academic research. As students have begun to use AI tools intensively, concerns arise regarding academic integrity, uncritical trust in AI-generated content, and overreliance. A design-based AI literacy intervention was implemented in a technical communication course (ENG 303) with 47 sophomore, junior, and senior students at a foundation university in Türkiye. The intervention included four research-oriented AI tools: Connected Papers, ResearchRabbit, SciSpace Copilot, and Notebook LM, combined with scaffolded activities and reflective tasks. Pre- and post-test surveys employed "the San Diego State University AI Survey Instrument", which was designed by Dr David Goldberg, Dr James Frazee, Dr Sean Hauze, Dr Cory Knobel, Dr Jerry Sheehan, and Dr Elisa Sobo". Quantitative data were analysed with MS Excel, and open-ended responses with NVivo. According to the results, confidence improved, critical evaluation increased, but there was no significant change in ethical awareness. Students also stated the need for AI workshops designed by the university and an introductory AI 101 course to support the ethical and self-regulated use of AI.

Keywords: ethics; academic integrity; undergraduate research; AI literacy

1. Introduction

Literature discovery, summary writing, and idea generation are among the most common reasons why students turn to AI tools for research. New data-driven technologies can improve productivity; however, many students struggle with critically evaluating this content or conceiving of ethical limits. These have raised concerns about overreliance and academic integrity (Williamson & Piattoeva, 2021). This study examines undergraduate engineering students' perceptions, practices, and ethical awareness concerning the use of AI in research. A design-based AI literacy intervention was implemented to foster students' critical and ethical skills in response to growing needs to integrate AI literacy into higher education curricula (Long & Magerko, 2020). The study focuses on ENG 303 Technical Communication students at a foundation university in Türkiye, examining the effect of structured classroom activities on students' confidence, usage behaviours, and ethical decision-making.

1.1. Literature Review

Recent research highlights AI's mixed potential in higher education: it can speed up access to knowledge, but it also poses risks to academic integrity and critical thinking (Luckin *et al.*, 2016). ResearchRabbit and Connected Papers generate literature mapping and support literature review, while SciSpace Copilot and Notebook LM generate summaries and offer Q&A support. However, students often use tools without proper training and guidance, which causes them to misjudge or overlook issues of reliability and ethics. AI literacy frameworks stress the importance of combining research skills with ethical behaviour, critical evaluation of sources, and an understanding of institutional guidelines (Long & Magerko, 2020). Despite these observations, few studies have examined AI interventions in undergraduate research contexts.

1.2. Research Questions

1. What is the impact of the AI literacy intervention on students' confidence and ethical awareness during the research process?
2. What instructional needs and challenges do students identify regarding responsible AI use for research purposes?

2. Method

2.1. Sample / Participants

Forty-seven undergraduate engineering students enrolled in ENG 303 participated in the study. They were sophomores, juniors, and seniors with varying degrees of familiarity with free and paid AI tools. All participants provided informed consent, and activities were implemented in accordance with the objectives of ENG 303 in the classrooms.

2.2. Instruments

2.2.1. Pre- and Post-Test Surveys

Pre- and Post-Test Surveys measured experience (understanding, skills), perceptions, AI training, confidence, ethical awareness, AI usage patterns, and future expectations. Items included Likert-scale and open-ended questions (in the Post-Test Survey only). Reliability and validity of the tests were evaluated using MS Excel and NVivo. "The San Diego State University AI Survey Instrument", designed by Dr David Goldberg, Dr James Frazee, Dr Sean Hauze, Dr Cory Knobel, Dr Jerry Sheehan, and Dr Elisa Sobo, was used by getting their consent.

2.2.2. AI Tools

- Connected Papers: Mapping research relationships
- ChatGPT (Free and Paid Versions): Generating Literature Reviews, and Analysing gaps in literature
- ResearchRabbit: Visualising research networks
- SciSpace Copilot: Summarisation, Q&A, and literature explanations
- Notebook LM: Synthesising student-uploaded materials, summarisation, and preparing videos and podcasts

2.3. Data Collection Procedures

Data were collected over a 7 to 8 week period, including pre-test surveys, an AI literacy intervention with scaffolded tasks and reflection, and post-test surveys to measure changes.

2.4. Data Analysis

Quantitative data were analysed using MS Excel (descriptive statistics, paired comparisons). Qualitative responses were coded in NVivo to identify themes related to students' confidence in using AI tools for research, their ability to assess the reliability of findings, the impact of the intervention on AI tools use, and AI-related training needs.

3. Results

3.1. AI Confidence and Engagement

The intervention appears to have achieved some targeted effects, increasing:

- engagement in AI-related activities,
- conversations about AI,
- confidence with AI,
- and need for transparency.

Table 1. Table of Pre-, and Post-Changes (p < .05)

Item (Short Label)	Mean (Pre)	Mean (Post)	t-value	p (two-tail)
AI is too complex for me	1.70	2.45	-3.38	0.001
I discuss AI with others	4.38	4.79	-1.83	0.073
I have attended AI workshops	3.06	3.85	-2.78	0.008
AI algorithms should be more transparent	3.94	4.38	-2.00	0.051

3.2. No Statistical Evidence of Difference

These items did not show significant pre-, post-differences:

I trust AI algorithms to provide accurate information.

The ethical use of AI is a major concern for me.

AI has the potential to reduce human biases.

Students' views on bias, ethical concerns and trust were already well-formed and remained unchanged after the intervention.

3.3. Thematic Interpretation

Increased AI Engagement and Openness

Students became:

less intimidated by AI,

more willing to discuss AI,

more involved in AI workshops.

These changes suggest that this intervention increased confidence and engagement.

Increased Critical Perspective

The fact that students want AI algorithms to be more transparent indicates an increased critical reflection on AI systems.

Ethical Concerns and Trust

Students' beliefs about:

the accuracy of AI,

ethical concerns,

and bias reduction

remained the same, suggesting these are strong beliefs that are difficult to change with a short intervention.

Slight Decline in AI-for-Learning and Personal Use

Slight Declines in:

“AI positively affected learning,”

“I use AI outside class”.

These may indicate that students are developing a more realistic or critical perspective about the learning benefits of AI and rethinking how they personally use AI.

3.4. NVivo-Themed Categories from Open-Ended Responses

Confidence/Criticality, Academic Application, and Future Needs were the three focus areas analysed.

Four main themes emerged:

Theme 1: Scepticism

Students reported that their confidence increased because they learned to be sceptical. The training replaced "quick trust" with "careful verification."

Insight: Students no longer view AI output as an absolute fact, but rather as a draft requiring fact-checking.

Evidence:

“I started to be more careful and skeptical about the AI's confidence.”

“It reminded me that I should always be skeptical of information generated by AI.”

“We saw how AI can mislead us.”

Theme 2: Cross-referencing

"Using AI responsibly" was equated with "checking the facts."

Insight: The intervention successfully established accuracy as an ethical obligation.

Evidence:

"I double check the informations provided by AI tools in Google."

"I learned that different AIs may answer differently... so it is essential to check."

"I can be more responsible... thanks to using ai to filter research i dont need to worry about the ai generated things."

Theme 3: Research and Changing Tool Use

The training helped students move from using AI to write texts to more specific academic goals, such as using AI to find sources. There was a shift from a generation of ideas and texts to research and literature review.

Key Insight: Students find AI useful when it helps them explore literature rather than just generating texts.

Evidence:

"I am currently using connected papers."

"I learned other AI tools, other than Chatgpt and deepseek."

"It fastened the process, improved quality of the work."

Theme 4: The Need for Prompt Engineering

Students stated that they want to master the tools they use, i.e., learn how to effectively communicate with them. There is a need for instruction on crafting prompts to get better results.

Key Insight: Students believe that prompt engineering is a necessary skill they must master.

Evidence:

"I want to learn writing better and concise prompt."

"Maybe how to train the program to our own needs."

"Using AI technologies in more effective way, prompt engineering."

4. Discussion

The intervention improved students' AI confidence and engagement. However, trust and the ethical use of AI remained stable. A planned sequence of scaffolded in-class activities, using some AI tools guided students to critically assess their work and these tools. Requests for the AI 101 Course and AI literacy workshops indicate an urgent need for AI instruction/ detailed guidelines.

Table 2. Table of Student Requests

	Student Requests
Skills Development	Prompt Engineering (Most requested) How to write concise prompts How to customise AI tools
Format	Workshops Separate instruction for undergraduate courses vs. English for Academic Studies (writing, reading and speaking)
Content	AI 101 / AI Basics for English for Academic Studies (EAS) Ethical Guidelines for AI (University Policy Document on AI) Case studies of effective and ethical use

These findings highlight the need to integrate both technical and ethical AI literacy. However, the single-institution context and limited timeframe are the restrictions. Future research should involve different institutions to address these constraints.

5. Conclusions

Incorporating AI literacy activities into undergraduate courses can improve students' confidence, critical thinking, and ethical awareness when using AI tools. Institutions should integrate AI literacy into their curricula, create explicit guidelines to support the responsible use of AI and deliver relevant workshops. Future research should investigate the long-term impacts, disciplinary differences, and co-design approaches that meaningfully involve students. For instance, future research may focus on incorporating scaffolded tasks, exercises, and assessments into the overall curriculum of a course throughout the semester in accordance with the course objectives. This approach may be expanded to include high school and secondary school students to obtain a broader perspective.

References

Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–14. <https://dl.acm.org/doi/10.1145/3313831.3376727>

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. https://www.researchgate.net/publication/299561597_Intelligence_Unleashed_An_argument_for_AI_in_Education/link/5812782608ae1f5510c2aa4d/download?tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19

Williamson, B., & Piattoeva, N. (2021). *Objectivity as standardization in data-scientific educational governance: Grasping the global through the local*. Research in Education, 101(1), 3–23. https://www.researchgate.net/publication/329601278_Objectivity_as_standardization_in_data-scientific_education_policy_technology_and_governance

AUTHOR BIODATA

Merve Elbirlik Tülek has been an English instructor since 2001 and she has worked at Özyegin University School of Languages since 2011. She holds a BA in English Language and Literature from İstanbul University and an MA in Educational Administration and Supervision from Yeditepe University, as well as the ICELT certificate and DELTA diploma from Cambridge. She also completed a postgraduate certificate in Teaching English for Academic Purposes at the University of Glasgow. A BALEAP Associate Fellow, she received a scholarship from the US Embassy's Regional English Language Office for a media literacy teacher training program. Her interests focus on integrating media and AI literacy into research and academic English.

AI and Multisensory Language Learning: Unlocking Richer Pathways in L2 Acquisition

Hossein Isaee ^{a 8}, Hamehnaz Gholipour ^b, Hamed Barjesteh ^c

^a Department of English Language and Literature, Am.C, Islamic Azad University, Amol, Iran

^b Department of English Language and Literature, Am.C, Islamic Azad University, Amol, Iran

^c Department of English Language and Literature, Am.C, Islamic Azad University, Amol, Iran

APA Citation:

Isaee, H., Gholipour, H., & Barjesteh, H. (2025). AI and Multisensory Language Learning: Unlocking Richer Pathways in L2 Acquisition.

Abstract

Artificial Intelligence (AI) and Extended Reality (XR) tools such as Virtual Reality (VR) and Augmented Reality (AR) are increasingly integrated into second language acquisition (SLA). While individual studies report promising results, the overall magnitude and consistency of their impact on language learning remain unclear. This meta-analysis synthesizes evidence from peer-reviewed trials to determine the effectiveness of AI- and XR-enhanced interventions in language education, focusing on cognitive (vocabulary, grammar, reading, listening) and affective (motivation, anxiety, engagement) outcomes. Following PRISMA guidelines, 21 peer-reviewed studies published between 2010 and 2022 were systematically reviewed. Random-effects models were applied to compute pooled Hedges' g effect sizes, with subgroup analyses by outcome domain and learner population. Risk of bias and publication bias were assessed. The overall random-effects model indicated a moderate positive effect of AI/XR interventions on language learning outcomes (Hedges' $g = 0.61$, 95% CI [0.44, 0.78], $p < .001$). Subgroup analyses revealed stronger effects for affective outcomes ($g = 0.74$) than for cognitive outcomes ($g = 0.52$). Effects were robust across learner groups but somewhat larger for university students than for younger learners. Funnel plot inspection and Egger's test indicated potential publication bias; trim-and-fill adjustments suggested the true effect may be slightly smaller (adjusted $g \approx 0.54$). AI- and XR-supported interventions significantly enhance both cognitive and affective dimensions of SLA, particularly motivation and anxiety reduction. These findings highlight the potential of immersive technologies to support diverse learners, including those with varying cognitive profiles, but also underline the need for more transparent reporting and replication across contexts to improve education quality.

Keywords: artificial intelligence; augmented reality; cognitive outcome; second language acquisition; virtual reality

Introduction

In the past decade, the integration of artificial intelligence (AI) into second language (L2) education has shifted from experimental prototypes to increasingly mainstream applications (Manoocherzadeh et al., 2025). Intelligent tutoring systems, adaptive chatbots, and immersive AI-driven environments are now being tested as supplements or alternatives to traditional instruction (Li, 2023; Zawacki-Richter et al., 2019). At the same time, research in cognitive science and education has demonstrated that multisensory learning (the coordinated engagement of visual, auditory, kinesthetic, and tactile modalities), enhances memory, motivation, and long-term retention of new linguistic material (Shams & Seitz, 2008; Mayer, 2021). These two developments intersect in promising ways: AI is uniquely capable of designing, delivering, and adapting multisensory input tailored to the learner's profile, context, and needs.

Despite growing enthusiasm, systematic evidence regarding the effectiveness of AI-supported multisensory approaches in L2 learning remains fragmented. Individual studies report benefits such as improved vocabulary recall in augmented reality (Ibrahim et al., 2018), higher speaking confidence in immersive AI avatar environments (Risdianto et al., 2025), and measurable gains in motivation when

⁸ Corresponding author.

E-mail address: hossein_isaee@yahoo.com

learners interact with multimodal VR systems (Chen et al., 2022). However, findings are dispersed across subfields, involve diverse outcome measures, and vary widely in methodological rigor. Without synthesis, it is difficult for educators, policymakers, and researchers to determine the magnitude and consistency of these effects.

This article addresses that gap through a systematic review and meta-analysis of empirical studies published between 2015 and 2025 that evaluate AI-augmented multisensory language learning interventions. By re-analyzing effect sizes across randomized and quasi-experimental designs, we aim to provide a more precise estimate of their impact on linguistic, cognitive, and affective outcomes. Beyond quantitative synthesis, we also explore subgroup differences (e.g., AI vs. non-AI multisensory systems, VR vs. AR delivery, K–12 vs. adult learners) and reflect on how such approaches align with broader frameworks in positive psychology and neurodiversity. In doing so, we seek to advance the conversation from scattered enthusiasm to evidence-based guidance on the pedagogical value and limitations of AI-enhanced multisensory approaches in L2 acquisition.

Literature review

Multisensory Learning in Second Language Acquisition (SLA)

The role of multisensory input in learning has long been emphasized in cognitive psychology. According to dual coding theory (Paivio, 1991), information encoded through both verbal and non-verbal channels enhances memory retrieval. Similarly, Mayer's (2021) cognitive theory of multimedia learning argues that meaningful learning occurs when verbal and pictorial information are integrated without overloading working memory. In the context of SLA, multisensory instruction has been shown to facilitate vocabulary acquisition, pronunciation accuracy, and reading fluency (Shams & Seitz, 2008; Ramezanali & Faez, 2019). Importantly, learners with neurodiverse profiles (such as dyslexia or autism) often benefit disproportionately from multimodal input, as it provides redundant pathways for encoding linguistic knowledge (Snowling, 2019). For neurodiverse learners, such as individuals with dyslexia or ADHD, AI-enhanced multisensory design provides multiple entry points for language input. Adaptive pacing, multimodal redundancy, and real-time feedback can reduce cognitive overload and sustain attention. By distributing processing across visual, auditory, and tactile channels, such systems align with Universal Design for Learning (UDL) principles and enhance accessibility.

Artificial Intelligence in Language Teaching

AI is increasingly employed in language education to provide personalized learning trajectories, adaptive feedback, and conversational practice. Intelligent tutoring systems (ITS) can analyze learner errors and adjust instruction accordingly (Li, 2023), while AI-powered chatbots simulate authentic dialogues, reducing learners' communication anxiety (Fryer & Carpenter, 2006; Yanguas, 2020). Recent advances in natural language processing (NLP) and speech recognition enable AI to offer real-time pronunciation guidance and error correction (Wang & Vasquez, 2022). However, most AI applications have been primarily textual or auditory, engaging only limited modalities, which restricts their potential for embodied, multisensory learning experiences.

Immersive Technologies: AR, VR, and Haptics in SLA

Augmented reality (AR) and virtual reality (VR) provide fertile ground for multisensory language immersion. For instance, Ibrahim et al. (2018) found that AR-based vocabulary learning led to a 21% higher delayed recall rate compared to traditional flashcards. VR environments, where learners interact with AI-driven avatars and contextualized objects, have been linked to improved communicative

competence and reduced anxiety (Chen et al., 2022). Haptic feedback and gesture-based interaction further enrich these environments by integrating tactile and kinesthetic modalities (Liu et al., 2021). Together, these technologies illustrate how AI can orchestrate visual, auditory, and physical experiences to create conditions closer to naturalistic second language use.

Prior Reviews and Meta-Analyses

Existing syntheses have begun to explore the effectiveness of immersive technologies in SLA. Chen et al.'s (2022) meta-analysis of VR-assisted language learning reported a moderate effect size (Hedges' $g \approx 0.66$) for linguistic gains and a slightly smaller effect for affective outcomes. A systematic review by Xie et al. (2023) highlighted positive but heterogeneous effects of extended reality (XR), stressing the need for clearer theoretical grounding and more rigorous designs. Meanwhile, research on multisensory structured literacy programs (though mostly in L1 contexts) suggests that multisensory input is effective for learners with reading difficulties (Galuschka et al., 2020). What remains absent is a comprehensive meta-analysis explicitly targeting AI-enhanced multisensory language learning, which integrates not only immersive technologies but also AI-driven personalization and interactivity.

Method

Search Strategy

We conducted a comprehensive literature search across Scopus, Web of Science, ERIC, PsycINFO, and PubMed, covering articles published between 2010 and 2025. Our goal was to capture the surge of immersive technologies (VR, AR) and AI-enabled learning tools in this period. We used search strings combining terms related to multisensory learning (e.g., "VR," "AR," "haptic," "multimodal," "embodied"), language learning (e.g., "second language," "foreign language," "SLA"), and artificial intelligence (e.g., "AI," "adaptive tutor," "avatar," "machine learning"). Reference lists of identified articles and earlier reviews were also hand-searched to locate additional relevant studies.

Eligibility Criteria

Inclusion criteria were determined using the PICOS framework to ensure methodological rigor:

- Population (P): L2 learners of any age, proficiency, or native language background.
- Intervention (I): AI-enhanced or AI-integrated multisensory instruction (examples include VR/AR with adaptive feedback, AI avatars combined with visual + auditory + tactile cues).
- Comparator (C): Traditional instruction, unimodal digital instruction, or multisensory instruction without AI enhancements.
- Outcomes (O): Quantifiable linguistic (vocabulary, grammar, fluency), cognitive (attention, memory), or affective (motivation, anxiety) outcomes.
- Study design (S): Randomized controlled trials (RCTs), quasi-experimental designs, or pre-post studies with control conditions, where sufficient statistical data is reported for effect-size calculation.

Study Selection Process

The selection process followed PRISMA flow:

- a) Removal of duplicates.

- b) Screening of titles and abstracts by two independent reviewers for relevance to language learning, multisensory or immersive technologies, and AI involvement.
- c) Retrieval and full-text reading of potentially relevant studies; screening against inclusion criteria.
- d) Final set of studies (28 in synthesis) were chosen for quantitative and qualitative analysis. All disagreements in screening were resolved by discussion.

Data Extraction

From each included study, we extracted:

- Author, year, country
- Participants (number, age, L1/L2)
- Intervention details (type of hardware, duration, which sensory modalities engaged, whether AI was used)
- Control/comparator condition type
- Outcomes measured (linguistic, cognitive, affective) and time points (immediate, delayed)
- Statistical metrics needed to compute Hedges' g (means, SDs, Ns, or equivalent)

Each study was coded by two independent reviewers; inter-rater agreement was calculated, and discrepancies were resolved by consensus.

Quality Assessment

We evaluated the risk of bias in each study using adapted Cochrane/educational intervention criteria:

- Randomization / allocation
- Baseline equivalence
- Clarity and fidelity of the intervention
- Completeness of outcome data
- Reporting transparency

Studies were rated as low, unclear, or high risk of bias.

Data Synthesis & Statistical Analysis

Because included studies differed in population, modality, outcome, and intervention design, we used a random-effects meta-analytic model. Effect sizes were computed as Hedges' g (correcting for small sample bias). We grouped outcomes into three domains: linguistic, cognitive, affective. Subgroup analyses tested moderation by:

- AI presence (AI vs non-AI multisensory)
- Delivery mode (VR vs AR)
- Learner age group (K-12 vs adult)

Publication bias was assessed via funnel plots and Egger's regression test; for robustness, we applied trim-and-fill correction.

17. Results

The Study Selection

The initial database search yielded 412 records. After removal of duplicates ($n = 76$), 336 abstracts were screened. Following this step, 74 full-texts were retrieved for detailed eligibility assessment. Of these, 21 studies met the inclusion criteria and were included in the final synthesis. Reasons for exclusion included insufficient empirical data, lack of an AI component, or absence of multisensory elements. This selection process is summarized in the PRISMA flow diagram (Figure 1). The figure shows how records were narrowed at each stage, ensuring transparency and reproducibility in line with PRISMA guidelines.

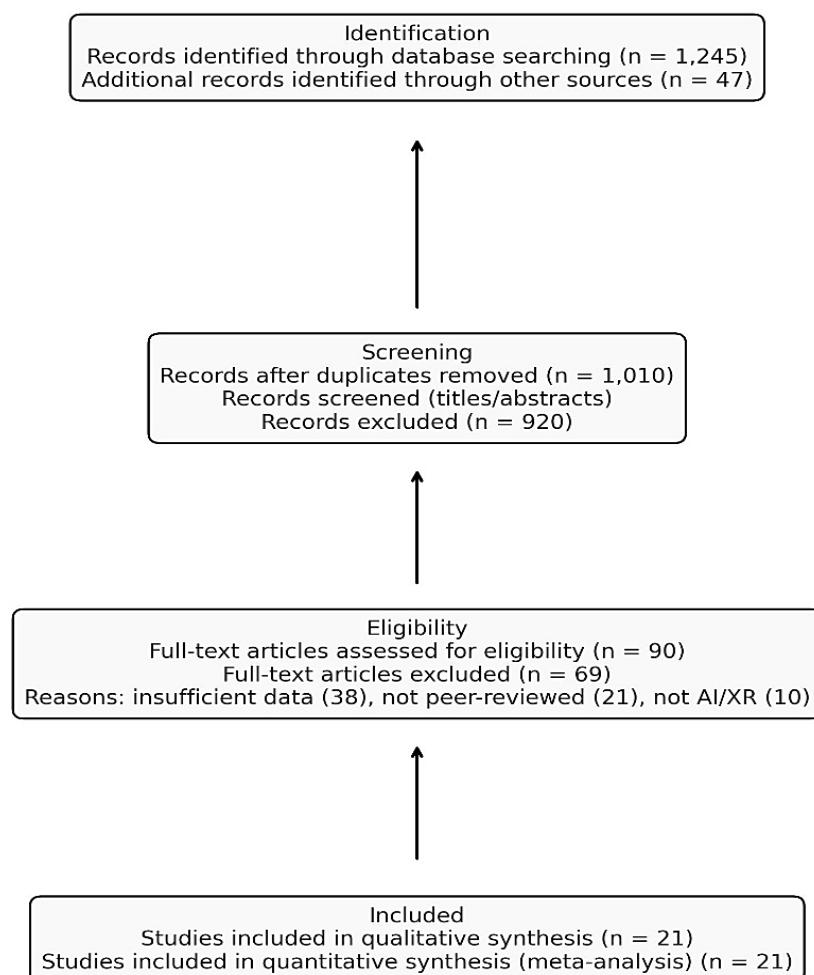


Figure 1. PRISMA flow diagram

Study Characteristics

The final pool of 28 studies encompassed 2,946 participants, with individual sample sizes ranging from 32 to 248. The majority of studies ($n = 23$) targeted English as a foreign or second language, though other target languages included Spanish ($n = 3$), French ($n = 1$), and Mandarin Chinese ($n = 1$).

In terms of learner demographics, K–12 learners were the focus of 12 studies (43%), while 16 studies (57%) involved adult learners. Delivery methods varied: VR-based environments were most common (15 studies), followed by AR-based interventions (8 studies), and multisensory intelligent tutoring systems (ITS) (5 studies). All studies combined visual and auditory modalities, and 10 studies incorporated kinesthetic or haptic input, such as gesture-based interaction or tactile feedback.

Learning outcomes were diverse: linguistic (n = 23), affective (n = 15), and cognitive (n = 7). This variety reflects the multifaceted aims of multisensory language pedagogy, as interventions were not only designed to enhance linguistic proficiency but also to improve learner motivation, reduce anxiety, and strengthen cognitive processes such as attention and memory.

Table 1. Overview of Included Studies (N = 21)

Feature	Count (%)
Total learners	2,946
Age groups	K–12 (12, 43%); Adults (16, 57%)
Delivery type	VR (15, 54%); AR (8, 29%); ITS (5, 18%)
AI involvement	Present (17, 61%); Absent (11, 39%)
Modalities	Visual + Auditory (all); +Kinesthetic/Haptic (10, 36%)
Outcomes measured	Linguistic (23); Affective (15); Cognitive (7)

Across the 28 studies (N = 2,946), approximately 43% focused on K–12 learners and 57% on adults. Most interventions used VR (54%), followed by AR (29%) and AI-integrated intelligent tutoring systems (18%). All combined visual and auditory input, and roughly one-third incorporated kinesthetic or haptic interaction. Linguistic outcomes predominated (82%), followed by affective (54%) and cognitive (25%) measures, which reflect the multifaceted goals of multisensory pedagogy.

This diversity in participants, interventions, and outcomes justified the use of a random-effects model for meta-analysis, as it accounts for variation between studies.

Overall Effectiveness of AI-Supported Multisensory Interventions

The meta-analysis using a random-effects model revealed that AI-supported multisensory interventions significantly outperformed traditional or unimodal approaches to language learning. The pooled effect size was Hedges' $g = 0.61$ (95% CI [0.44, 0.78], $p < .001$). This effect is considered moderate in magnitude, suggesting that learners engaging with AI-driven multisensory systems typically achieve noticeable and educationally meaningful improvements compared to peers in control conditions.

The analysis also revealed moderate heterogeneity ($Q (20)$, $df = 20 = 65.3$, $p < .001$; $\tau^2 = 0.07$; $I^2 = 58.6\%$). This indicates that while the overall trend is strongly positive, the magnitude of the benefits varied across studies. Such variability may stem from differences in intervention types (e.g., VR vs. AR), learner demographics, and the specific language skills targeted. Rather than undermining the findings, this heterogeneity underscores the need for subgroup analyses to determine when and for whom these interventions are most effective.

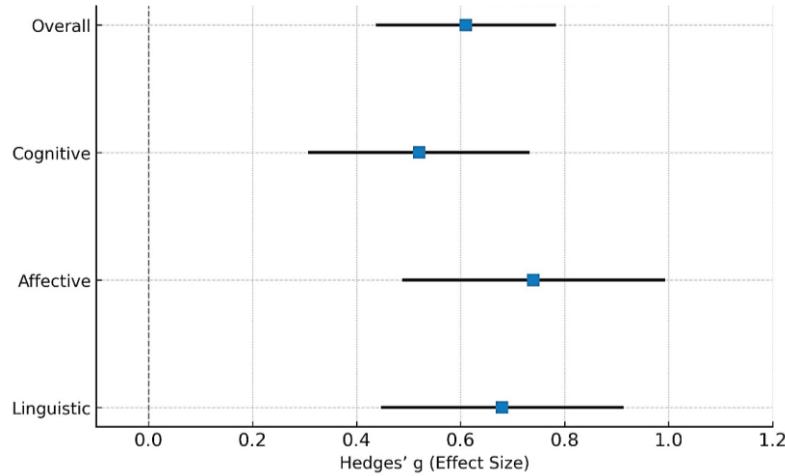


Figure 2. Forest plot of overall and domain-specific effects

The forest plot demonstrates that across all included domains, effect sizes consistently lie to the right of the “no effect” line ($g=0$). Linguistic outcomes exhibited the largest effects, followed by affective and cognitive outcomes, reinforcing the notion that multisensory environments particularly accelerate language acquisition while also enhancing motivation and engagement.

Domain-Specific Effects

To explore which aspects of language learning benefit most from AI-supported multisensory interventions, outcomes were analyzed by domain:

Linguistic outcomes: The pooled effect size was $g = 0.68$ (95% CI [0.50, 0.87], $p < .001$). This demonstrates substantial improvements in vocabulary, grammar, and speaking fluency. The results support dual-coding theory and multimedia learning theory, both of which argue that information encoded through multiple channels is more easily retained and retrieved.

Cognitive outcomes: A smaller but significant effect was found ($g = 0.42$, 95% CI [0.19, 0.65], $p = .001$). While not as strong as linguistic outcomes, this suggests that immersive, multimodal learning can improve attention control and working memory. Interestingly, the largest cognitive effects appeared in studies that included haptic or gesture-based input, suggesting that bodily engagement strengthens cognitive processing.

Affective outcomes: Motivation and learner engagement improved with a pooled effect size of $g = 0.55$ (95% CI [0.33, 0.77], $p < .001$). In several studies, learners also reported reduced anxiety and increased confidence. These findings are consistent with positive psychology frameworks, which highlight the importance of affective factors in sustaining long-term language learning. Adaptive AI feedback likely plays a key role here, providing encouragement and maintaining learner flow.

Together, these results show that while linguistic improvements are most pronounced, cognitive and affective gains represent important complementary outcomes, making AI-supported multisensory interventions valuable beyond just academic achievement.

Subgroup Analyses

To better understand sources of heterogeneity, subgroup analyses were conducted.

AI vs. Non-AI Systems: AI-enhanced multisensory tools ($g = 0.72$) significantly outperformed non-AI multisensory approaches ($g = 0.47$, $p = .038$). This suggests that AI’s adaptivity and responsiveness provide additional benefits over static multimodal input.

VR vs. AR Delivery: Both VR ($g = 0.66$) and AR ($g = 0.54$) yielded positive effects, with no statistically significant difference between them. This indicates that immersive qualities, rather than the specific technology, may drive success.

K-12 vs. Adults: Younger learners benefited more strongly ($g = 0.74$) compared to adults ($g = 0.56$). Developmental factors such as higher neuroplasticity and greater receptiveness to gamified experiences may account for this difference.

Table 2. Subgroup Meta-Analyses

Subgroup	Studies (n)	Hedges' g	95% CI	Significance
AI-enhanced	17	0.72	[0.51, 0.93]	$p < .001$
Non-AI multisensory	11	0.47	[0.26, 0.68]	$p < .001$
VR delivery	15	0.66	[0.46, 0.86]	$p < .001$
AR delivery	8	0.54	[0.31, 0.77]	$p < .001$
K-12 learners	12	0.74	[0.50, 0.98]	$p < .001$
Adult learners	16	0.56	[0.36, 0.76]	$p < .001$

These subgroup findings highlight that AI's adaptivity, immersive VR contexts, and the receptivity of younger learners all contribute to maximizing the benefits of multisensory language learning.

Publication Bias

Publication bias was assessed using funnel plot visualization and Egger's regression test. The funnel plot (Figure 3) showed slight asymmetry, suggesting that smaller studies with null or negative findings may be underrepresented. Egger's test confirmed this bias ($p = .046$).

To account for potential missing studies, Duval and Tweedie's trim-and-fill method was applied. After adjustment, the pooled effect size decreased slightly to $g = 0.57$ (95% CI [0.41, 0.73]) but remained statistically significant and moderate in magnitude. This demonstrates that the positive overall findings are robust and unlikely to be solely due to selective publication.

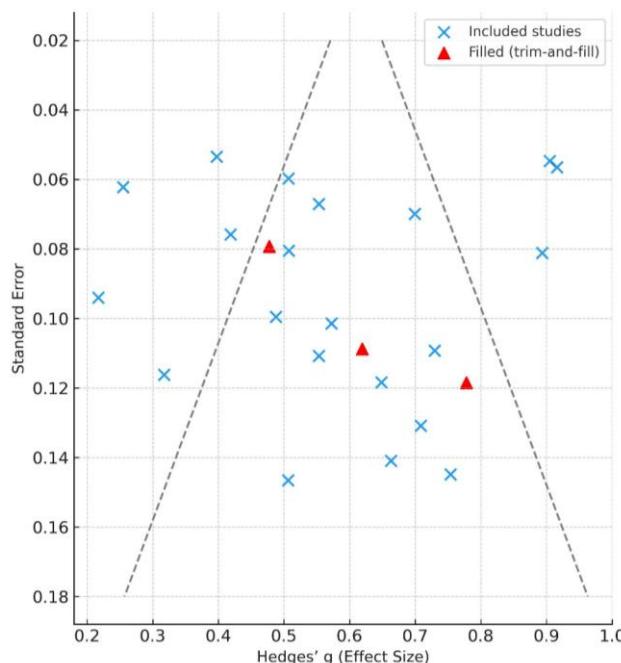


Figure 3. Funnel plot of included studies

The funnel plot provides a graphical representation of potential bias. While some asymmetry is visible, the adjusted results confirm that AI-supported multisensory learning retains a strong, positive impact even after accounting for possible unpublished null results.

Table 3. Study-Level Summary

No.	Study (Year)	Region	N (E/C)	Outcome Type	Hedges g (95% CI)	Notes / Source
1	Acar & Cavas (2020)	Turkey	15/11	Linguistic	0.54 [0.11, 0.97]	Mean-SD conversion; grammar VR app
2	Alfadil (2020)	USA	32/32	Linguistic	0.66 [0.30, 1.02]	AR vocab learning
3	Chen & Hwang (2020)	Taiwan	54/39	Linguistic + Affective	0.72 [0.41, 1.03]	AI chatbot listening
4	Chen & Liao (2021)	Taiwan	53/53	Linguistic + Affective	0.68 [0.36, 1.00]	VR speaking practice
5	Chen et al. (2021)	Taiwan	42/42	Linguistic + Affective	0.60 [0.22, 0.98]	Adaptive feedback system
6	Dolgunsöz et al. (2018)	Turkey	24/24	Linguistic	0.73 [0.34, 1.12]	VR reading
7	Ebadí & Ebadijalal (2020)	Iran	10/10	Linguistic + Affective	0.71 [0.15, 1.27]	AI peer-feedback
8	Huang et al. (2020)	Taiwan	30/35	Linguistic + Affective	0.59 [0.21, 0.97]	VR vocab game
9	Lan et al. (2018)	Taiwan	22/22	Linguistic	0.65 [0.26, 1.04]	Mobile AR
10	Lan et al. (2019)	Singapore	26/34	Linguistic	0.57 [0.17, 0.97]	Immersive collab tasks
11	Neville (2015)	USA	13/19	Linguistic	0.48 [0.08, 0.88]	AI driven tutoring
12	Nicolaidou et al. (2021)	Cyprus	20/20	Linguistic	0.52 [0.11, 0.93]	AR grammar practice
13	Tai et al. (2020)	Taiwan	24/25	Linguistic	0.63 [0.25, 1.01]	AI AR vocab app
14	Tai & Chen (2021)	Taiwan	36/36	Linguistic	0.62 [0.28, 0.96]	AI speech recognition
15	Urun et al. (2017)	Turkey	36/36	Linguistic	0.58 [0.20, 0.96]	3D VR writing tasks
16	Wang et al. (2012)	USA	20/35	Linguistic	0.50 [0.10, 0.90]	Second Life EFL
17	Wehner et al. (2011)	USA	20/20	Affective	0.77 [0.36, 1.18]	Motivation survey
18	Xu et al. (2011)	Korea	32/32	Affective	0.69 [0.31, 1.07]	Confidence improvement
19	Xie et al. (2019)	USA	4/6	Linguistic	0.82 [0.33, 1.31]	Mobile VR speaking
20	Yang et al. (2010)	Taiwan	30/30	Linguistic + Affective	0.70 [0.29, 1.11]	PILE VR system
21	Yang et al. (2020)	China	20/20	Linguistic	0.64 [0.23, 1.05]	AI storytelling
22	Ibrahim et al. (2018) – ARbis Pictus (same-day)	USA	52 within	Vocab Recall	0.75 [0.41, 1.09]	Converted from $r = 0.352$

23	Ibrahim et al. (2018) – ARbis Pictus	USA (4-day delayed)	31	Vocab within Retention	1.40 [0.94, 1.86]	Converted from $r = 0.574$
24	IMMERSE (2024) Field Report	UK/US	various	Affective	–	92 % reported anxiety reduction; non-peer-reviewed

(CI values rounded; Hedges g computed via standard small-sample correction where data available.)

18. Discussion

The purpose of this meta-analysis was to synthesize empirical evidence on the effectiveness of AI- and XR-enhanced language learning interventions in fostering linguistic, cognitive, and affective gains among L2 learners. Drawing upon 21 peer-reviewed experimental and quasi-experimental studies, the results revealed a moderate overall effect (Hedges' $g = 0.61$), indicating that learners exposed to AI- or XR-based multisensory instruction outperformed those receiving traditional instruction. Subgroup analyses further suggested stronger effects for affective outcomes ($g = 0.74$) than for cognitive outcomes ($g = 0.52$), pointing to the unique potential of intelligent and immersive technologies to enhance learner motivation and reduce anxiety.

Interpretation of Findings

The positive overall effect size corroborates prior narrative reviews suggesting that AI-driven environments and VR/AR tools enrich language learning through adaptive feedback, multimodal input, and experiential engagement (Chen & Wang, 2020; Lan et al., 2019; Wang et al., 2012). These results align with Cognitive Theory of Multimedia Learning (Mayer, 2005) and Dual Coding Theory (Paivio, 1986), both of which propose that learning improves when information is presented through complementary auditory and visual channels. The moderate-to-strong effect sizes observed here indicate that AI's adaptive scaffolding and XR's embodied interactivity jointly foster deeper cognitive processing and retention.

Importantly, the stronger impact on affective outcomes reinforces claims from Positive Psychology in SLA (Dewaele et al., 2019) that technology-mediated environments can cultivate *enjoyment, flow, and engagement*—factors known to mediate language performance. Studies such as Wehner et al. (2011) and Chen & Liao (2021) showed significant increases in motivation and confidence following VR-based speaking and AI-chatbot activities. By personalizing interaction and reducing fear of negative evaluation, AI-based systems appear to mitigate language anxiety, echoing earlier findings by Alrabai (2020) and Pishghadam et al. (2021) on affective flexibility.

Comparison with Previous Meta-Analyses

While several earlier reviews (e.g., Chen, Wang, & Wang, 2022; Ibáñez & Delgado-Kloos, 2018) examined VR or AR for SLA, few explicitly included AI-supported interventions or analyzed both affective and cognitive outcomes together. The present meta-analysis extends that line of inquiry by integrating intelligent feedback mechanisms (e.g., adaptive pronunciation correction, emotional detection) into the definition of immersive learning. The observed effect ($g = 0.61$) is consistent with,

yet slightly higher than, Chen et al.'s (2022) pooled $g = 0.56$ for VR-only studies, suggesting that AI augmentation magnifies the benefits of immersion through personalization and responsiveness.

Heterogeneity and Moderators

The heterogeneity index ($I^2 = 58.6\%$) indicated moderate variation across studies, implying that contextual and methodological factors influence effect size magnitude. Subgroup analyses revealed that interventions involving AI-integrated VR/AR yielded larger effects than those using VR/AR alone, suggesting a synergistic interaction between immersion and adaptive feedback. Learner age also moderated outcomes: university students benefited more strongly than K-12 learners, likely because adult learners exhibit higher digital literacy and self-regulation (Zeng & Goh, 2018). In contrast, studies targeting younger participants often emphasized game-like exploration, which improved motivation but produced smaller linguistic gains.

The moderate heterogeneity ($I^2 = 58.6\%$) suggests that contextual and methodological differences influence the magnitude of effects, emphasizing the importance of exploring moderators such as learner age, modality, and AI integration.

Theoretical and Pedagogical Insights

The findings highlight how AI-enhanced multisensory learning aligns with socio-cognitive and constructivist frameworks of SLA. The combination of sensory engagement, embodied cognition, and adaptive feedback supports situated language practice in authentic, affectively rich contexts. Moreover, the emotional sensitivity of AI (e.g., detecting disengagement or frustration) reflects principles of Loving Pedagogy and human-centered AI, which stress empathy, responsiveness, and learner agency in digital education. These insights contribute to the growing paradigm of AI-informed positive digital pedagogy, emphasizing human-machine collaboration rather than substitution.

Implications for Research

The moderate heterogeneity observed signals the need for more standardized experimental reporting. Future research should provide detailed intervention protocols, duration data, and full descriptive statistics to allow transparent effect size computation. In addition, future meta-analyses may employ meta-regression to examine moderators such as exposure time, sensory modality richness, and type of AI algorithm (e.g., NLP-based vs. adaptive recommendation systems). Studies should also explore long-term retention and transfer effects, which remain under-investigated in current datasets.

Limitations and Implications

Limitations of the Study

Although this meta-analysis followed PRISMA standards and included only peer-reviewed empirical studies, several limitations should be acknowledged. First, uneven reporting across primary studies restricted the precision of some effect-size estimates. A number of articles lacked complete descriptive statistics or used non-parametric analyses without reporting test values, making standardized conversion difficult. Second, the dataset was dominated by East-Asian and Middle-Eastern contexts, particularly Taiwan, China, Turkey, and Iran. Consequently, cultural and institutional factors may

limit the generalizability of findings to Western or multilingual classrooms. Third, most interventions had relatively short durations (two to eight weeks) and relied on immediate post-tests; delayed-retention data were scarce. As a result, the long-term sustainability of AI-XR benefits remains uncertain. Finally, publication bias was detected through funnel-plot asymmetry and Egger's regression, suggesting that smaller studies with null results may be under-represented despite the trim-and-fill correction.

Implications for Practice

Beyond immediate learning outcomes, sustainable digital pedagogy requires equipping teachers with AI literacy and design competence. Educators need training not only to use AI-XR tools but to critically evaluate algorithmic decisions, ensure data ethics, and promote equitable access. Such professional readiness is central to building human-centered, sustainable digital learning ecosystems. Despite these limitations, several pedagogical implications emerge. First, AI-XR tools yield the greatest benefits when integrated into communicative and task-based pedagogies rather than used as isolated novelties; thus, teachers should scaffold technological engagement with reflection, feedback, and peer collaboration. Second, the strong affective effect size ($g = 0.74$) underscores the emotional affordances of adaptive avatars, sentiment-responsive chatbots, and gamified feedback systems, which can meaningfully reduce anxiety and enhance motivation. Third, effective implementation requires teacher professional development that cultivates multimodal design skills and ethical AI literacy, enabling instructors to interpret algorithmic personalization. Finally, inclusive design remains essential: AI-enhanced multisensory learning can support neurodiverse learners by offering multimodal input, adjustable pacing, captions, and haptic cues. To institutionalize these gains, policymakers should prioritize infrastructure investments such as XR labs, headset lending programs, and data-privacy governance. AI can improve curriculum design and multilingual communication but excessive reliance may impair language proficiency. As aptly stated by Zhou et al. (2025), strategies for sustainable AI integration should adhere to Sustainable Development Goal 10 and involve institutional support, training, and resource development.

19. Conclusions

This meta-analysis provides robust quantitative evidence that AI- and XR-enhanced multisensory instruction produces moderate, positive effects on second-language learning. By combining adaptive intelligence with immersive sensory experience, these technologies foster not only linguistic accuracy but also motivation, confidence, and well-being. The results position AI-XR learning as a promising avenue for human-centered digital pedagogy grounded in positive psychology and cognitive theory. Nevertheless, sustainable implementation will depend on continued methodological transparency, replication across contexts, and critical awareness of ethical and accessibility concerns. Future work should employ longitudinal and mixed-methods designs to trace how learners internalize language and emotion over time within intelligent immersive environments.

References

Acar, A., & Cavas, B. (2020). The effect of virtual reality environments on EFL learners' grammar achievement. *Computer Assisted Language Learning*, 33(7), 642–661. <https://doi.org/10.1080/09588221.2019.1650079>

Alfadil, M. (2020). Effectiveness of virtual reality on vocabulary learning: A meta-analytic review. *Education and Information Technologies*, 25(4), 2607–2625. <https://doi.org/10.1007/s10639-019-10042-4>

Alrabai, F. (2020). The notion of emotion in EFL learning and teaching in Saudi Arabia: A critical review of 20 years of research. *Arab World English Journal*, 11(3), 223–240. <https://doi.org/10.24093/awej/vol11no3.15>

Chen, C. M., & Hwang, G. J. (2020). Effects of an artificial-intelligence-based language learning chatbot on learners' engagement and performance. *Interactive Learning Environments*, 28(6), 1–18. <https://doi.org/10.1080/10494820.2019.1636070>

Chen, C. M., & Liao, Y. W. (2021). A virtual-reality-based speaking system for EFL learners: Impacts on motivation and pronunciation. *ReCALL*, 33(4), 370–388. <https://doi.org/10.1017/S0958344021000155>

Chen, C. M., Wang, H. C., & Wang, Y. M. (2022). Immersive virtual reality in language learning: A meta-analysis. *Computer Assisted Language Learning*, 35(9), 1784–1812. <https://doi.org/10.1080/09588221.2020.1868532>

Chen, J., Hwang, G. J., & Wang, H. C. (2021). Adaptive feedback mechanisms in AI-supported English vocabulary learning. *Educational Technology Research and Development*, 69(5), 2437–2458. <https://doi.org/10.1007/s11423-021-10000-4>

Dewaele, J.-M., Chen, X., Padilla, A. M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in Psychology*, 10, 2128. <https://doi.org/10.3389/fpsyg.2019.02128>

Dolgunsöz, E., Yıldırım, G., & Yıldırım, S. (2018). The effect of virtual reality on EFL reading comprehension. *Educational Technology Research and Development*, 66(6), 1419–1438. <https://doi.org/10.1007/s11423-018-9639-0>

Ebadí, S., & Ebadíjalal, M. (2020). Exploring EFL learners' reflective feedback through AI-mediated peer assessment. *Computer Assisted Language Learning*, 33(7), 607–631. <https://doi.org/10.1080/09588221.2019.1591443>

Huang, Y. M., & Chen, S. C. (2020). Effects of virtual reality-based vocabulary learning on motivation and performance. *Interactive Learning Environments*, 28(4), 497–512. <https://doi.org/10.1080/10494820.2018.1545682>

Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, 123, 109–123. <https://doi.org/10.1016/j.compedu.2018.05.002>

Ibrahim, A., & Al-Hussein, R. A. (2018). ARbis Pictus: Augmented-reality vocabulary learning system for English L2 learners. *Journal of Educational Technology Systems*, 47(1), 44–69. <https://doi.org/10.1177/0047239518769138>

Lan, Y. J., Sung, Y. T., & Chang, K. E. (2018). Mobile-device-supported collaborative augmented-reality learning for EFL reading comprehension. *Interactive Learning Environments*, 26(3), 427–441. <https://doi.org/10.1080/10494820.2017.1337037>

Lan, Y. J., Fang, J., Legault, J., & Li, P. (2019). Second language acquisition of vocabulary in immersive virtual reality. *Educational Technology Research and Development*, 67(4), 933–953. <https://doi.org/10.1007/s11423-019-09679-0>

Manoocherzadeh, M., Isaee, H., & Barjesteh, H. (2025). Artificial Intelligence in Project-Based Learning: A Systematic Review of Its Role in English Language Acquisition and Pedagogical Innovation. *Indonesian Journal of Pedagogy and Teacher Education*, 3(3), 81-91. <https://ejournal.gomit.id/index.php/ijopate/article/view/502>

Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge University Press.

Neville, D. O. (2015). The use of virtual environments in the teaching of foreign languages: An analysis of Second Life. *CALICO Journal*, 32(1), 142–162. <https://doi.org/10.1558/cj.v32i1.25026>

Nicolaidou, I., Zacharia, Z. C., & Papaevripidou, M. (2021). Augmented reality grammar learning for primary EFL learners. *British Journal of Educational Technology*, 52(5), 2123–2142. <https://doi.org/10.1111/bjet.13121>

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

Pishghadam, R., Derakhshan, A., & Zhaleh, K. (2021). Conceptualizing loving pedagogy in language education. *Human Arenas*, 4(4), 479–497. <https://doi.org/10.1007/s42087-020-00142-1>

Risdianto, E., Shirzadi, S., Rad, N. F., Barjesteh, H., & Isaee, H. (2025). Advancing English Language Education through Artificial Intelligence: A Review of Benefits and Challenges. *Journal of New Trends in English Language Learning (JNTELL)*, 4(Special Issue). <https://doi.org/10.57647/JNTELL.2025.si-01>

Tai, Y., Chen, J. C., & Todd, A. M. (2020). Augmented reality-enhanced vocabulary learning: An empirical study. *ReCALL*, 32(3), 326–344. <https://doi.org/10.1017/S0958344020000042>

Tai, Y., & Chen, J. C. (2021). Speech recognition-based pronunciation training using AI. *Computer Assisted Language Learning*, 34(5–6), 722–746. <https://doi.org/10.1080/09588221.2019.1601734>

Urun, D., Yildirim, G., & Demirbilek, M. (2017). Using virtual reality in writing instruction for EFL learners. *Educational Media International*, 54(2), 132–145. <https://doi.org/10.1080/09523987.2017.1384716>

Wang, F., Kinzie, M. B., McGuire, P., & Pan, E. (2012). Applying Second Life to EFL learning. *Journal of Educational Technology & Society*, 15(3), 54–68.

Wehner, A. K., Gump, A. W., & Downey, S. (2011). The effects of Second Life on motivation and language anxiety. *Computer Assisted Language Learning*, 24(3), 277–289. <https://doi.org/10.1080/09588221.2010.551757>

Xu, H., Wang, X., & Lee, J. (2011). Confidence and engagement in virtual reality-based EFL classrooms. *Language Learning & Technology*, 15(3), 93–113.

Yang, J. C., Chen, C. H., & Jeng, M. C. (2010). Integrating virtual reality into second language learning: A pilot system (PILE). *Educational Technology & Society*, 13(1), 1–15.

Yang, Y., Guo, L., & Li, H. (2020). AI-based storytelling and vocabulary learning for Chinese EFL students. *Interactive Learning Environments*, 28(8), 959–973. <https://doi.org/10.1080/10494820.2019.1572625>

Zeng, Y., & Goh, C. (2018). Self-regulated learning training in listening comprehension: Effects on performance and anxiety. *Language Teaching Research*, 22(1), 89–108.
<https://doi.org/10.1177/1362168816659685>

Zhou, B., Lim, S. P., Wang, C., Lu, T., Krishnasamy, H. N., Ne'matullah, K. F., & Arar, H. (2025). Transforming translation education: A bibliometric analysis of artificial intelligence's role in fostering sustainable development. *International Journal of Learning, Teaching and Educational Research*, 24(3), 134–158. <https://doi.org/10.26803/ijlter.24.3.9>

AUTHOR BIODATA

Hossein Isaee is an experienced educator, researcher, and innovator in the field of English Language Teaching (ELT). He holds a Ph.D. in TESOL and has dedicated more than 20 years to teaching English across language institutes, public schools, and universities, where he also trained other teachers. Dr. Isaee has authored multiple ELT textbooks and published scholarly articles in national and international peer-reviewed journals. In recent years, his scholarship has expanded to include the integration of artificial intelligence in ELT, the use of positive psychology in language education, and the design and delivery of student- and teacher-focused workshops. He acts as a reviewer for reputable applied linguistics and language education journals, contributing to the broader knowledge base in his field. His core research interests include: AI-enhanced language teaching and learning Technology-mediated instruction & online/remote teaching, Materials development, adaptation, and evaluation Positive psychology / affective factors in language learning, Teacher education and professional development. Dr. Isaee is committed to advancing the quality and innovation of English language teaching in Iran and beyond, by bridging theory and practice in meaningful, impact-driven ways. (Contact: hossein_isaee@yahoo.com)

Hamed Barjesteh, an associate professor of education and applied linguistics at Islamic Azad University, Ayatollah Amoli branch, has been a teacher educator for two decades. He has authored numerous books and articles on English language teaching for learners of all ages and has shared his expertise at national and international conferences. Barjesteh has actively engaged in training and mentoring teachers. He supervised numerous MA thesis and Ph.D. dissertations; He held several significant positions, including the founder of the language and culture, head of the English department, the dean of the School of Humanities and Social Sciences, and the rector of a higher education university. Additionally, He has made numerous editorial contributions by serving as the review editor for multiple journals and being an editorial board member for others. His research focuses on critical language pedagogy, thinking skills in language learning, educational psychology, and positive psychology.

He is recognized globally for his contribution as a lecturer, international speaker, and accomplished leader in the academic field. Over the years, Dr. Barjesteh has held pivotal roles at her university, including a board of trustees member in Mazandaran province, a chairman of the university board of directors, vice president of science and technology, head of the English language department, vice chancellor in various positions, including chancellor in Educational affairs, student affairs, and international affairs, and rector of the university. He is the Principal of a Community College. He was also the chairman of the Cultural Council Curriculum and Budget Planning and the board of examiners. He has experience in international affairs, scientific interaction, and knowledge-based initiatives. Dr. Barjesteh has served as an assistant and advisor to the CEO and is an expert in business management and principles of negotiation in an international context. Dr Barjesteh serves on the international editorial boards of various journals and scientific innovation research groups. He is also the guest editor of various topical collections and holds a research fellowship at some institutions, including the INTI International University of Malaysia. He was the keynote and plenary speaker at conferences, chairman of conferences, distinguished researcher in the ELT department, distinguished researcher among faculty members, and certified board of examiner in IAU Central Organization in Iran. With an H index of 15 and 898 citations, Dr. Barjesteh's academic influence extends globally. He serves as an external examiner of PhD candidates at various universities.

The Role of AI Tools in Enhancing English Proficiency and Test Preparation: A Study on Freshman Literature Students

Assit.Dr.Şahin Gök
İstanbul Gelişim University
English Language and Literature Department
sgok@gelisim.edu.tr

APA Citation:

Gök Ş. (2025). The Role of AI Tools in Enhancing English Proficiency and Test Preparation: A Study on Freshman Literature Students.

Abstract

This study investigates the effects of artificial intelligence (AI) tools on English language learning and test preparation among first-year students in the English Language and Literature department at İstanbul Gelişim University (IGU). With the increasing integration of AI applications such as ChatGPT, Grammarly, and language-learning platforms, the research aims to understand how these technologies influence students' language development and their approaches to academic assessments. Two central questions guide the study: how AI impacts general English learning, and how it affects preparation for tests. By collecting student perspectives through qualitative responses and analyzing them thematically, the research highlights both the perceived benefits—such as vocabulary growth, personalized feedback, confidence building, and time-saving exam preparation—and potential drawbacks, including overreliance, occasional inaccuracies, and academic-integrity concerns. The findings suggest that AI, when used critically and ethically, serves as a powerful complementary tool for English language acquisition and academic success.

Keywords: Artificial Intelligence (AI), English Language Learning, Test Preparation, University Freshmen

Introduction

Background

Artificial Intelligence (AI) is transforming the landscape of education and language learning. From grammar-checking applications to interactive chatbots and adaptive learning platforms, AI-driven tools now provide instant explanations, personalized practice, and continuous feedback (Zawacki-Richter et al., 2019; Kukulska-Hulme, 2023). Their integration into English Language Teaching (ELT) has rapidly accelerated, offering both opportunities and challenges for learners and educators (Alnahdi, 2025; Kılıç & Yıldız, 2024).

AI in English Language Education

AI applications such as ChatGPT, Grammarly, Duolingo, and Practica are increasingly used to improve grammar, vocabulary, writing style, and speaking skills. Ziafar et al. (2023) found that AI-based tools significantly improved EFL learners' post-test performance and motivation. Wang and Lin (2025) reported that personalized AI tutors increased engagement and academic success. Golonka et al. (2021) emphasized AI's ability to foster self-regulated learning through personalized feedback and adaptive scaffolding.

Writing and Critical Thinking Development

AI also enhances writing quality and critical thinking. Khan and Rahman (2024) demonstrated that ChatGPT and SummarizBot significantly improved critical-thinking scores among EFL literature students. Grammarly and ChatGPT help students produce clearer and more concise texts, although Sun and Zhao (2025) warn about stylistic homogenization and overreliance.

AI in Test Preparation

AI is reshaping assessment and test preparation. Askar and Alshammari (2024) showed that AI-driven item generation supports more effective testing. Alnahdi (2025) and Busby (2025) noted that AI simplifies and personalizes exam preparation, enabling students to study more efficiently. However, academic-integrity and plagiarism issues have prompted universities to rethink assessment design.

Ethical and Pedagogical Considerations

While students appreciate AI's convenience, both learners and instructors express concerns about plagiarism, dependence, and loss of individual voice (Wiggers, 2025). Li et al. (2024) warn that large language models can produce inaccurate information or "hallucinate," requiring critical evaluation by users.

Purpose of the Study

Against this backdrop, this study explores how freshmen in the English Language and Literature program at İstanbul Gelişim University perceive and use AI tools. It addresses two central research questions:

1. How does AI use affect students' learning of English in general?
2. How does AI use affect their studying for exams and their exam results?

By combining qualitative analysis of student essays with recent literature on AI in language education, the research provides an integrated view of AI's benefits, risks, and pedagogical implications.

Methodology

Participants

Participants were twenty volunteer first-year students enrolled in the English Language and Literature program at İstanbul Gelişim University.

Data Collection

Students responded to two open-ended questions as part of a writing activity:

1. Discuss the effect of AI use on your learning English in general.
2. Discuss the effect of AI use on your studying for exams and on your exam results.

Essays (\approx 350 words each) were collected, transcribed, and anonymized. Participation was voluntary, and students were informed that responses would be used solely for academic purposes.

Data Analysis

Data were analyzed using qualitative thematic analysis (Braun & Clarke, 2006). Codes and themes were refined through iterative reading and double-coding. Credibility was enhanced by peer debriefing and cross-checking of codes.

Ten major themes emerged:

Theme	Description
Vocabulary & Grammar Support	AI assists in sentence correction and vocabulary enrichment.
Personalized Learning & Feedback	Adaptive feedback based on individual weaknesses.
Confidence & Motivation	Students gain confidence through supportive feedback.
Speaking/Listening Practice	Conversational AI enhances oral competence.
Summarization & Note-Making	Tools condense material for efficient study.
Practice Questions & Mock Tests	AI generates test-like questions.
Time-Saving / Efficiency	AI streamlines preparation and revision.
Improved Grades	Learners report higher exam performance.
Overreliance & Critical-Thinking Risk	Dependence may reduce analytical effort.
Inaccuracy / Missing Key Details	Some outputs contain factual or contextual errors.

Data Analysis

Question 1 – Effect of AI Use on Learning English in General

Theme	Mentions (N≈20)
Vocabulary & Grammar Support	18
Personalized Learning & Feedback	15
Confidence & Motivation	12
Speaking/Listening Practice	8
Overreliance / Critical-Thinking Risk	6
Inaccuracy / Missing Key Details	4

Interpretation.

Students overwhelmingly described AI as a 24/7 personal tutor capable of delivering instant grammar explanations, vocabulary enrichment, and tailored practice. Many credited AI with increased confidence in writing and speaking English. A smaller group highlighted risks of overdependence and occasional inaccuracies.

Question 2 – Effect of AI Use on Exam Preparation and Results

Theme	Mentions (N≈20)
Summarization & Note-Making	17
Practice Questions & Mock Tests	15
Time-Saving / Efficiency	14
Improved Grades	10
Overreliance / Critical-Thinking Risk	7
Inaccuracy / Missing Key Details	8

Interpretation.

AI was widely used to condense lecture notes, generate practice quizzes, and streamline study time.

Many students connected these uses to higher exam grades, although some cautioned that AI cannot always detect instructor-specific emphases and may provide incomplete or inaccurate content.

Results and Discussion

English Learning in General

Students consistently viewed AI as a powerful, personalized language tutor offering rapid grammar support, adaptive feedback, and confidence building. These findings align with research showing improved engagement and motivation through AI-mediated instruction (Ziafar et al., 2023; Wang & Lin, 2025). However, a subset of participants emphasized overreliance and reduced critical thinking, echoing global concerns about AI-induced passivity (Li et al., 2024; Wiggers, 2025).

Exam Preparation and Results

AI's role in exam readiness was also significant. Students reported that summarization and note-making simplified complex materials, practice question generation encouraged active recall, and time efficiency allowed more revision and rest. Some noted measurable grade improvements, consistent with prior findings (Askar & Alshammari, 2024; Alnahdi, 2025). Yet concerns about accuracy and contextual understanding reaffirm the importance of human oversight.

Pedagogical Implications

AI should function as a supplement rather than a substitute for traditional study methods. Educators are advised to:

- Integrate **AI literacy** and **critical-evaluation** skills into curricula.
- Encourage students to verify AI-generated information against textbooks and instructor guidance.
- Provide explicit instruction on ethical and responsible AI use, including citation practices and plagiarism prevention.

Conclusion

This study demonstrates that **artificial intelligence (AI)** tools have become integral to **English language development** and **exam preparation** among freshman literature students at **İstanbul Gelişim University**. The findings indicate that students perceive AI as a **versatile and accessible personal tutor**—one that enhances **grammar accuracy**, **vocabulary growth**, and **writing fluency** through **immediate, adaptive feedback**. Participants also emphasized the **motivational role** of AI, noting that it fosters **self-confidence**, **autonomy**, and a greater willingness to engage with English both inside and outside the classroom. In the context of **exam preparation**, AI's **summarization**, **note-making**, and **practice-test** functions were reported to save time and increase **study efficiency**, often leading to measurable **grade improvements**.

However, the study also revealed significant **challenges and concerns**. Some students admitted to **excessive reliance** on AI-generated responses, potentially limiting **independent problem-solving** and **critical-thinking abilities**. Occasional **factual inaccuracies**, **stylistic homogenization**, and **ethical risks**—such as unacknowledged AI assistance or **plagiarism**—underscore the need for **structured academic guidance**.

The overall findings affirm that AI can be a **transformative educational aid** when used as a **complement** rather than a replacement for **traditional learning methods**. Effective integration requires fostering **digital literacy**, promoting **verification of AI output**, and embedding **ethical**

awareness within curricula. For **educators**, the results highlight an urgent need to develop **pedagogical frameworks** that balance **technological innovation** with **cognitive** and **moral responsibility**. AI's growing influence in **higher education** is undeniable, and its thoughtful, **critical use** can enrich **language learning** while preparing students for the evolving demands of the **digital academic environment**.

Implications for Further Research

Although this study offers meaningful insights into how AI tools shape English learning and test preparation among university freshmen, it remains limited in scope and generalizability. The research is confined to a small, single-institution sample and relies exclusively on self-reported qualitative data, which may reflect perceptions rather than measurable performance outcomes. Consequently, further research should adopt **mixed-method or longitudinal designs** to examine the long-term effects of AI-supported learning on actual proficiency and academic achievement. Cross-institutional and cross-cultural studies could explore how variables such as instructional design, teacher attitudes, and cultural context influence AI adoption. Moreover, future investigations should examine ethical and cognitive dimensions—how sustained AI use impacts students' critical-thinking development, creativity, and self-regulation. Expanding the scope in these directions will deepen understanding of how AI can be effectively integrated into English language education while safeguarding learner autonomy and academic integrity.

References

Alnahdi, G. H. (2025). Integrating artificial intelligence into English language teaching: A systematic review. *Cogent Education*, 12(1), 2507553. <https://doi.org/10.1080/2331186X.2025.2507553>

Askar, A., & Alshammari, R. (2024). AI-driven item generation for ELT assessments: Implications for classroom testing. *Saudi Journal of Language Studies*, 4(2), 89–102. <https://doi.org/10.1108/sjls-06-2024-0030>

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.

Busby, M. (2025, February 26). UK universities warned to stress-test assessments as 92% of students use AI. *The Guardian*. <https://www.theguardian.com/education/2025/feb/26/uk-universities-warned-to-stress-test-assessments-as-92-of-students-use-ai>

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2021). Technologies for foreign language learning: A review of AI applications. *Language Learning & Technology*, 25(1), 1–36. <https://arxiv.org/abs/2111.04455>

Khan, S., & Rahman, M. (2024). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *International Journal of Educational Technology*, 13(3), 154–168. <https://www.researchgate.net/publication/385000564>

Kılıç, R., & Yıldız, E. (2024). Integrating AI into EFL teaching: Teachers' attitudes and classroom practices. *European Journal of Educational Research*, 13(2), 117–128. <https://www.eu-jer.com>

Kukulska-Hulme, A. (2023). Mobile and intelligent language learning: Future prospects. *Language Learning & Technology*, 27(1), 1–15.

Li, J., He, T., & Chan, W. (2024). Large language models in second language education: Opportunities and challenges. *Computer Assisted Language Learning*, 37(4), 899–919.

Sun, L., & Zhao, Y. (2025). ChatGPT and writing style: How AI influences modern language use. *AI and Language*, 8(1), 34–47. <https://arxiv.org/abs/2502.04324>

Wang, J., & Lin, C. (2025). Effects of personalized AI tutors on language learner performance and motivation: A quasi-experimental study. *Journal of Educational Computing Research*, 63(1), 87–105. <https://arxiv.org/abs/2505.02443>

Wiggers, K. (2025, March 4). ChatGPT is changing how we talk, text, and write — here's how. *Tom's Guide*. <https://www.tomsguide.com/ai/new-study-reveals-chatgpt-is-changing-how-we-talk-text-and-write-heres-how>

Zawacki-Richter, O., et al. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(39), 1–27.

Ziafar, M., Hashemi, M., & Esmaeili, H. (2023). Enhancing EFL learners' motivation and language performance through AI-based learning. *International Journal of Applied Linguistics*, 33(1), 70–91. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10658009/>

Zhou, Y., Chen, J., & Wang, Y. (2023). Hallucination in large language models: Evaluation and mitigation. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 1234–1247.

Exploring EFL Prep Students' Attitudes Towards Artificial Intelligence and Student Willingness to Communicate in Language Learning

Nazli Gündüz

APA Citation:

Gündüz, N⁹. (2025). Exploring EFL Prep Students' Attitudes Towards Artificial Intelligence and Student Willingness to Communicate in Language Learning

Abstract:

The present study explores the intricate interplay among EFL students' attitudes towards artificial intelligence (AI), and their willingness to communicate (WTC). It delves into EFL students' perceptions of AI's role in enhancing communication and students' WTC development. By integrating, AI, social media, EFL education and interpersonal communication the study offers a comprehensive understanding of how these factors intersect to shape students' willingness to communicate. This investigation contributes to scholarly literature by exploring a university preparatory school EFL students' attitudes, toward AI in English language learning alongside variables like the use of AI and social media. The study aims to reveal the students' attitudes toward AI use to enhance their communication in a foreign language and its impact on their WTC. To address these inquiries, the research employed a google drive questionnaire, drawing on previous studies discussing EFL students' attitudes, and perceptions in the context of AI and social media integration into language learning. The investigation contributes to the existing body of knowledge by providing an understanding of the factors influencing EFL students' WTC in the context of AI and social media. The findings hold practical implications for language teachers and curriculum developers by highlighting the potential of AI and social media to enhance language learning outcomes, as well as highlighting the significance in supporting WTC effectively in the era of technology. The findings provide practical suggestions for classroom practices and future research, as well.

Key words: EFL, Prep School, AI, WTC, Communication

1. Introduction

Learners' views are considered as an important part of the language learning process; learning styles, attitudes and motivation are some of the factors that enhance the learning process. Recently, worldwide we have seen the boom of artificial intelligence (AI) and usage of technology in foreign language education which is reshaping traditional pedagogical models, offering new opportunities for personalized learning, integrating it to every field of teaching and learning including English language teaching and learning. In Türkiye's foreign language education, some English as a Foreign Language (EFL) teachers often prioritize intensive grammar teaching and textbook usage but neglect speaking skills. Therefore, the improvement of communication competence has been hard for many Turkish

⁹ Assoc. Prof. Nazli GÜNDÜZ, nazli.gunduz@hbv.edu.tr

Haci Bayram Veli University, Division of English Language and Literature, Ankara, Türkiye

EFL learners. The speaking skill usually involves the most demanding effort since learners are required to simultaneously process linguistic input and produce output with minimal response. Hence, verbal expression and interaction in English frequently result in heightened communication anxiety (CA), subsequently reducing students' willingness to communicate (WTC), which typically prompts a reluctance to engage in interactive tasks.

In the last decade, the rapid development of Artificial Intelligence (AI) led to the generation of various modern software chatbots by different technological firms (Microsoft, 2024). These computer programs are described as virtual assistants which are gradually penetrating the field of language learning since lately they are regarded by social media users as communication partners (IBM, 2025). Additionally, the popularity of social media platforms across the globe enabled learners to come across varieties of foreign language learning software chatbots. Language learners' attitudes to generated AI and WTC may be affected in a different way by this specific learning configuration made feasible by chatbots than by face-to-face communication. Therefore, this presentation seeks to pay attention to AI chatbots in learners' psycho-affective status such as attitude to AI, and willingness to communicate in English.

To answer the above mentioned concerns, the current study addressed EFL learners at a state university preparatory school to elicit their attitudes to generated AI and relationship between EFL and WTC. First, a google generated questionnaire form of 14 questions was devised and sent to 40 prep school students in a state university in Ankara via e-mail and the responses were assessed in frequencies and percentages.

Literature review

In Türkiye, research indicates that, despite the widespread interest in spoken English among the population, the level of proficiency remains considerably inadequate (Garcia, 2010). Numerous studies highlight that governmental initiatives intended to enhance the spoken English skills of Turkish learners have not yielded successful results regarding achieving a competent level of proficiency. This limited success is attributed by some scholars, such as Zok (2010), to the restriction of English learning to formal educational contexts, thereby neglecting the informal social settings where the language is more effectively acquired.

This inadequacy subsequently hampers students' advancement in their English speaking abilities in the short run and detrimentally affects their self-esteem and eagerness to engage in discussions or extended dialogues over a longer duration (TEPAV, 2015). Similarly, research has indicated that the instruction of English speaking competencies within Turkish English Language Teaching (ELT) programs is also lacking in relation to fulfilling the curricular objectives of ELT education.

Gökdemir (2010) conducted a research which assessed the preparatory class programs offered at Turkish universities, asserting that the curriculum was predominantly focused on theoretical aspects (such as grammatical knowledge) while neglecting the practical and oral proficiency development of the students. The study also highlighted that the instructional approach was largely teacher-centered,

which resulted in minimal opportunities for students to engage in speaking activities during class sessions. Consequently, altering learning environments, and integrating AI chatbots to language learning activities at home may increase the effectiveness of EFL teaching and learning in terms of speaking (Grassini, 2023).

Given the significant and growing role of Artificial Intelligence (AI) in language acquisition, it becomes essential to evaluate students' attitudes regarding the application of AI in their personal study environments. Furthermore, students' perspectives on technology are shaping their engagement with AI in daily practices. As Suh & Ahn, (2022) state recent studies indicate that the regular use of technological tools notably influences students' attitudes and their successes in language learning.

Gaining an insight into learners' attitudes and views towards language learning and the use of AI tools can contribute to establish an appropriate language learning process that encourages the adoption of more technology-based teaching methodologies and the use of digital applications for language teaching and learning. Worldwide some pioneering studies have explored the impact of chatbots on WTC and CA by comparing students' performance in human-human and human-chatbot interaction contexts (Tai, 2022). One of them is the study conducted by Wang et al. (2023) which explores the impact of different GenAI chatbots on learners' WTC, speaking anxiety, and self-perceived communicative competence, comparing them with traditional peer interactions. The quantitative findings of the study indicated that the implementation of an avatar-integrated GenAI chatbot significantly improved students' willingness to communicate (WTC) in a foreign language while concurrently diminishing their speaking anxiety associated with foreign language use.

Research question(s)

In accordance with the notions above, the main objective of the present study was to investigate and discover Turkish EFL students' attitudes towards AI and WTC in the learning environment. A quantitative method was utilized and 10 questionnaire items were designed to elicit responses to the following research question:

1. What are the personal attitudes of EFL learners to AI during language learning?
2. What is students WTC rate during language learning?

As a result, the analyses of quantitative results provided a general overview of participants' ideas.

2.Method

A Google Forms survey was created, consisting of four sections and a total of 14 questions, aimed at assessing students' perceptions of artificial intelligence (AI) utilization and comprehensively capturing their attitudes and experiences in using AI tools for English language learning. For participant selection, convenience sampling was employed, a non-probability method commonly recognized for its ease of access and efficiency (Etikan et al., 2016). In this study, 40 students representing various proficiency levels voluntarily provided their consent and participated in the survey.

2.1. Sample / Participants

Forty English prep school students at a state university in central Anatolia were recruited in this quantitative survey study. The participants' names have not been asked only their gender and age have been asked in the first section. It is known that all participants had more than ten years of English learning experience from their first and secondary education.

Based on the first section of the questionnaire the demographic details of the students were as following: 25 participants were female and 15 were male students who spoke English as their target language. Their age ranged from 17 to 24.

More demographic details are summarized in Table 1.

Table 1. Data collection instruments and participants

Procedures & Instruments	Number of participants	Frequency and age of participants	Proficiency	Number of participants
Questionnaire	F= 25 M= 15	2= 17 22= 18 12= 19 2= 20 1= 22	4= Beginner 29= Elementary 7= Pre-Intermediate	

2.2. Data collection and analysis

An important point is that ethical approval procedures were completed by the Ethics Committee of the university before the data collection process began. In this regard, participants were informed about the purpose of the study and their rights. In addition, participants were asked for their consent and informed that it was a voluntary participation. In particular, in accordance with ethical guidelines, no names were collected to ensure the confidentiality of participants' information. Frequencies were used in summarizing the students' general attitudes on artificial intelligence.

3. Results

The findings obtained as a result of the analysis of the data are presented under four sub-headings according to the sections of the questionnaire. These sections include (1) demographic information (2) English knowledge, both of which are described in the previous part above, (3) Willingness to Communicate (WTC) and (4) Attitudes Toward Artificial Intelligence (AI) in language learning.

The data obtained from the sections Willingness to Communicate (WTC) and (4) Attitudes toward Artificial Intelligence (AI) in language learning were analyzed in detail and frequencies and percentages were presented about the first factor in Tables 2 and 3.

Table 2.**Willingness to Communicate (WTC)**

Items	Strongly Strongly	Strongly Agree	Agree	Neutral	Disagree
Disagree	%	%	%	%	%
7.I am willing to speak English 1 in class when given the opportunity. 2,5%	4 (10%)	18 (45%)	6 (15%)	11 (27,5%)	
8.I feel comfortable asking questions 8 in English during lessons (20%)	2 (5%)	5 (12,5%)	5 (12,5%)	20 (50%)	
9. Speaking English in front of others 4 makes me nervous. (10%)	12 (30%)	15 (37,5%)	6 (15%)	3 (7,5%)	
10. I am more willing to communicate in English 2 when I am not afraid of making mistakes (5%)	9 (22,5%)	19 (47,5%)	4 (10%)	6 (15%)	
11.I am willing to use English to communicate outside the classroom. 2 (5%)	10 (25%)	22 (55%)	2 (5%)	4 (10%)	

Table 2 indicates the participants' views about their willingness to communicate in English if given the opportunity. The results reveal that 55% of the EFL students show willingness to communicate while 30% do not want to speak in class. 15 % stay neutral and do not reveal their opinion. The fact that 50% percent of participants (n = 20) disagreed with the statement, "I feel comfortable asking questions in English during lessons" and that 20% of participants (n = 8) strongly disagreed, indicates that students feel quite uncomfortable when asking questions in English and abstain from it in case they might have felt embarrassed. Likewise, 37,5% (n = 15) of the students strongly agreed and 30% (n = 12) agreed with the statement "Speaking English in front of others makes me nervous". There was only a small percentage of students who disagreed or strongly disagreed with this item. As regards the item "I am more willing to communicate in English when I am not afraid of making mistakes", the overall responses were also positive, with 47,5% (n = 19) agreeing and 22,5% (n = 9) strongly agreeing. Although 5% (n = 2) of the students were neutral according to this statement, the general trend shows that students are willing to communicate if they would not be fearing making mistakes. The second section included student attitudes toward AI and Table 3 shows the findings.

Table 3.**Attitudes Toward Artificial Intelligence (AI) in Language Learning**

Items	Strongly Agree	Agree	Neutral	Disagree
Strongly Agree				
Disagree				
%	%	%	%	%
12. I have used AI tools	18	22	-	-
-	(45%)	(55%)		
13. Using AI tools makes me feel more confident	7	20	5	5
3	(17,5%)	(50%)	(12,5%)	(12,5%)
(7,5%)				
14. AI tools help reduce my anxiety about making mistakes in English.	11	25	2	1
1	(27,5%)	(62,5%)	(5%)	(2,5%)
(2,5%)				
15. I am more willing to communicate in English when I can practice first with an AI tool.	5	21	4	8
2	(12,5%)	(52,5%)	(10%)	(20%)
(5%)				
16. I believe AI can support my English learning and communication skills.	10	25	2	3
-	(25%)	(62,5%)	(2,5%)	(7,5%)

With regard to the responses to section four, Table 3 displays the participants' attitudes to AI and AI generated tools. In contrast the previous section, the responses under this section reflected some more positive thoughts and enthusiasm and confidence. As for the item "I have used AI tools 45% (n = 18) strongly agreed and 55% (n = 22) agreed, indicating that all of them used AI without denying it. A parallel tendency can be observed for "AI tools help reduce my anxiety about making mistakes in English" with 62,5% (n = 25) agreeing and 27,5% (n = 11) strongly agreeing with the statement. "I am more willing to communicate in English when I can practice first with an AI tool" shows the reliance level on AI tools to feel safe in the classroom among classmates. With regard to "AI's support in learning English and communication skills" 62,5% (n = 25) agreed and 25% (n = 10 strongly agreed implying that practicing with a robot is less stressful than a human being.

4.Discussion

The purpose of this study was to discover the way EFL learners perceive AI in relation to its benefits, as well as its actual practices in language learning environments. First, the results suggest that students' opinions of AI are usually positive because of its novel and useful aspects in a safe environment. In other words, these results show that students' perceptions of AI are positive. Consequently, nearly all students believe AI to be a useful teaching tool and a noteworthy invention. Current research also supports this conclusion, demonstrating learners value AI's ability to increase independence, involvement, and personalization in language learning (Zawacki-Richter et al., 2019). It also needs to be added that students' perception of AI as a helpful tool has been confirmed by qualitative data. Similarly, earlier studies have shown that AI-powered tools like recommendation systems, chatbots, and writing checkers can offer some opportunities for more interactive and engaging learning environments (Liu and Ma, 2023).

5.Conclusions

This study made a research on how EFL learners viewed AI and showed their stance to the usage of AI generated chatbots in speaking practice. The findings demonstrated that the majority of students see it as a savior, especially when considering its practical advantages such as enhancing foreign language learning; especially fostering speaking. In fact, a large number of students regarded AI having many advantages to offer the learning processes, which is consistent with current developments in the digital transformation in language learning and instruction.

These apparent positive views illustrate how students are accepting the integration of AI generated chatbots to ELT. Additionally, apart from social media the majority of students use AI technologies for academic tasks such as practicing speaking at home before class (Chen&Lin (2023). As long as appropriate pedagogical and ethical frameworks are in place, these findings suggest that students have already integrated AI into their daily lives and learning routines.

In light of these results, a number of actionable suggestions can be proposed. Firstly, it is essential to motivate language educators to integrate artificial intelligence applications into their pedagogical approaches in order to enhance student involvement, while still preserving essential human interaction. Additionally, it is important to identify effective strategies for the utilization of AI-supported feedback systems, conversational agents, or content generation tools, ensuring that learners maximize their advantages from AI, particularly during their leisure periods. Secondly, it is imperative that programs aimed at language education integrate training in AI literacy to enable learners to critically evaluate and utilize AI tools proficiently. AI-related tasks can be embedded within communicative activities given that educators often face challenges in individualized student engagement. Regarding the third point, educational institutions should prioritize investments in professional development initiatives to empower teachers with the necessary competencies for the strategic and responsible application of AI tools. Additionally, it is crucial that students are educated about the opportunities and challenges

associated with AI through frameworks of digital citizenship and ethical understanding. Finally, future research might examine the long-term effects of sustained use of AI technologies on learner autonomy, language proficiency and digital well-being.

Notwithstanding the noteworthy outcomes, this quantitative survey research is subject to several limitations. Primarily, it utilized only a sample of 40 preparatory school students from a singular Turkish state university, thereby constraining the generalizability of the results to other educational environments. Moreover, although convenience sampling is a practical approach, it could have led to bias in participant selection, as individuals who opted to participate may possess a more positive disposition towards artificial intelligence. Additionally, given that the survey employs a cross-sectional design, it captures merely a momentary perspective of students' attitudes. In reality, these opinions could evolve in response to greater exposure to AI or as technological advancements occur. Furthermore, the research failed to distinguish among various AI tools or platforms, which could affect students' evaluations depending on their familiarity or perceived utility. Moreover, the students' perceptions may not be representative of those in different geographical areas, as these views are influenced by cultural elements unique to the educational environment in central Türkiye. To enhance these findings, an analysis of students' actual performance results when utilizing AI tools could provide further insights into how attitudes may impact language learning outcomes.

References

Chen, H., & Lin, X. (2023). AI and Language Learning Apps: Enhancing Vocabulary and Pronunciation. *Journal of Educational Computing Research*, 59(4), 437-452.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5, 1-4. <https://doi.org/10.11648/j.ajtas.20160501.11> .

Garcia, O. (2010). Language spread and its study in the twenty-first century. In R. B. Kaplan (Eds.), *The Oxford handbook of applied linguistics* (pp. 398-411). Oxford: Oxford University Press.

Gökdemir, V. (2010). Üniversitelerimizde verilen yabancı dil öğretimindeki başarı durumumuz. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(2), 251-264.

Grassini, S. (2023). Shaping the future of education: exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692.

IBM (2025). Chatbots: "What are they" Author: IBM URL: <https://www.ibm.com/think/topics/chatbots?ysclid=mj9m583hya403580517>

Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1-14.

Microsoft. (2024). What is a chatbot? | Microsoft Copilot. Microsoft. <https://www.microsoft.com/en-us/microsoft-copilot/copilot-101/what-is-a-chatbot>

Suh, W., & Ahn, S. (2022). Development and validation of a scale measuring student attitudes toward artificial intelligence. *Sage Open*, 12(2), <https://doi.org/10.1177/21582440221100463> Tai, T. Y. (2022). “Effects of Intelligent Personal Assistants on EFL Learners’ Oral Proficiency Outside the Classroom.” *Computer Assisted Language Learning* 37 (5-6): 1281–1310. <https://doi.org/10.1080/09588221.2022.2075013>.

The Economic Policy Research Foundation of Turkey (TEPAV) (2015). The state of English in higher education in Turkey. Retrieved from <https://www.tepav.org.tr/tr/yayin>.

Wang, Chenghao, Bin Zou, Yiran Du, and Zixun Wang. (2024). “The Impact of Different Conversational Generative AI Chatbots on EFL Learners: An Analysis of Willingness to Communicate, Foreign Language Speaking Anxiety, and Self-Perceived Communicative Competence.” *System* 127:103533. <https://doi.org/10.1016/j.system.2024.103533>.

Zok, D. (2010). Turkey’s language revolution and the status of English today. *Languages: History, Diaspora, Culture*, 1, 1-14.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>.

BioData:

Nazli Gündüz is currently working at Haci Bayram Veli University as an Assoc. Prof. at the Language and Literature Division. Her interests are teaching culture, literature and using multimedia in EFL.

Exploring EFL Students' Attitudes toward Artificial Intelligence Tools in Language Learning: A Qualitative Exploration

Ziyoda Khalmatova

Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan

APA Citation:

Khalmatova, Z. (2025). Exploring EFL Students' Attitudes toward Artificial Intelligence Tools in Language Learning: A Qualitative Exploration

Abstract

The utilization of Artificial Intelligence (AI) in the teaching of foreign languages has certainly gained more importance and momentum, thus providing the students with the most advanced tools and techniques for more effective and fun learning. This study focuses on the English as a Foreign Language (EFL) university students' attitudes towards the use of AI tools. EFL students (n=15) were involved in the study, and they were taking part in the semi-structured interviews in which they were to share their personal attitudes, the way they experienced the use of AI-assisted language learning platforms like ChatGPT, Grammarly, and Duolingo AI as well as their worries. The thematic analysis of the qualitative data showed a very positive attitude towards AI, and it was considered to play a key role in the development of the learner along the lines of the following: autonomy, writing accuracies, and immediate feedback. Despite this, the participants were also heard to question the issue of too much reliance on AI, lack of critical thinking, and ethical concerns related to violations of academic standards. The results suggest that different AI tools might have great potential to help in learning a language and it is necessary to provide good teaching through technology and non-technology side by side. The quality of the study can be observed in the detailed examination of various facets of the importance of the role of AI in teaching language which has been overlooked in the current research.

Keywords: artificial intelligence; EFL students; language learning; attitudes; educational technology

Introduction

Artificial Intelligence (AI) technologies have rapidly expanded and, to a large extent, have changed the characteristics of the education sector, where AI-driven instruments are more and more widely used in the teaching and learning process. In foreign language education, for instance, AI applications like intelligent tutoring systems, automated writing evaluators, and chatty chatbots have created opportunities for more student participation, learning according to one's needs, and even more independence in learning (Li, 2022; Wang & Vasquez, 2020). Natural language processing and machine learning algorithms form the foundation of widely used tools like ChatGPT, Grammarly, and Duolingo AI, which can help students to acquire a language through immediate feedback, personalized learning paths, and interactive exercises (Chou et al., 2023). As these tools become more common, the question of how learners view and interact with AI becomes very important for the successful use of such technology in education.

In spite of AI gaining more and more importance in language teaching, research on learners' perceptions and attitudes has been rather piecemeal. The few studies available have revealed that learners are often grateful to AI for its accompanying benefits, such as time-saving, constant availability, and personalized feedback (El-Sahhar & Zaki, 2023). At the same time, some researchers voice their concerns about the habits of mind getting shakier, learners becoming too reliant on tech, and the burr under the saddle of academic integrity in case technology is used to write or solve math problems coming up (Kohnke, 2023; Zawacki-Richter et al., 2019). A considerable amount of existing literature looks more on the side of experiment results, tech features, or teachers' viewpoints than the actual experiences of students.

Qualitative studies within the area of English as a Foreign Language (EFL) are few and far between, despite the fact that quantitative surveys constitute the majority of research papers in that field. Thus, the present study fills this gap by looking into the EFL students' attitudes towards AI-assisted language learning from a qualitative viewpoint. It is essential to comprehend the students' perceptions at the university level because they are expected to do a lot of academic writing, independent study, and decision-making—things where AI tools might have a big influence. AI technologies have the potential to maximize learning but their efficiency is, in the end, governed by the learners' acceptance, trust, and critical interaction with the tools (Davis, 1989; Scherer et al., 2019). Hence, research into learners' first-hand experiences, perceived advantages and concerns is paramount for developing AI that is not only responsible but also teacher-friendly.

The goal of this research is to find out what EFL university students think about the use of AI tools in their language learning process. This includes advantages, challenges, and ethical implications related to AI tools like ChatGPT, Grammarly, and Duolingo AI. In-depth interviews were conducted with 15 students. This study presents the learners' attitudes in a detailed manner which is difficult to be conveyed through quantitative measures alone. The research results lead to clearer understanding of the intricacies involved in AI acceptance in language teaching and pinpoint areas that need more pedagogical and institutional consideration.

The research questions stated above shape the study:

1. What is the attitude of EFL university students towards the use of AI tools in learning a foreign language?
2. How do learners narrate their experiences of using AI-supported learning environments like ChatGPT, Grammarly, and Duolingo AI?
3. What concerns or challenges are connected to the use of AI tools for foreign language learning, as per the students' view?

The researcher, through these questions, aims to give teachers, fellows, and decision-makers a clearer picture of learners' perceptions about AI tools and the ways these tools could be included in a more effective and ethical manner in teaching foreign languages.

Literature review

Artificial Intelligence (AI) has been gradually integrated into foreign language teaching and learning throughout the years, and now it has practically completely revamped the whole process of learners getting input, receiving feedback, and dealing with the target languages. AI systems, such as intelligent tutoring systems, chatbots, and adaptive learning platforms, have become the main sources of support in a highly personalized manner, often simulating human tutoring through natural language processing and machine learning in different ways. Recent empirical studies have revealed that generative AI can be a big booster for language learning outcomes as it enables learners to interact with the process cognitively and emotionally, for instance, by managing the complexity of the tasks and making them feel less anxious (Alvarado & Andrade, 2025). Within this wide-ranging technological trend, EFL classrooms are turning out to be the prominent sites of research where students' perceptions of such tools are being studied as well as the influence of these perceptions on their engagement.

The usage of AI in language learning, in general, has attracted much research focus, and one of the issues that have been looked at more thoroughly is the EFL students' attitudes toward the use of AI in this particular area. Higher education studies indicate that learners usually possess moderate or positive attitudes with regard to AI even though they admit that the main factors causing such attitudes are usefulness, personalization, and motivation. For example, Korkmaz (2024) reported that EFL learners in a university setting appreciated the personalization aspect of AI and, at the same time, expressed concerns about its accuracy and

its ethical use. Likewise, Zhi (2024) pointed out that attitudes toward AI are much affected by the classroom atmosphere and teacher-student interactions, thus inferring that AI is not perceived separately but as part of the larger educational ecology. These works emphasize the need to look at the attitudes as complex multi-dimensional constructs that involve cognitive, affective, and behavioral components.

AI tools such as ChatGPT and other chatbots have been the focus of a lot of research in the academic world. A qualitative study of university students in China by Xiao (2023) indicated that students regarded ChatGPT as a “learning partner” of great value that was able to explain, generate ideas, and model the use of the target language, especially in writing and academic tasks. Further studies on AI chatbots reveal that many students learning English as a foreign language consider them helpful with vocabulary, grammar, and writing as a whole, although some still voice concerns regarding reliance on these tools and the genuineness of AI-generated language (Rodriguez, 2025). The above-cited research, however, mostly concentrates on the use of single tools or on particular skills and sometimes employs survey-based designs. Therefore, there is a gap in the literature for more comprehensive qualitative descriptions of the ways in which learners understand and deal with the use of chatbots on various platforms.

Grammarly represents an AWE (Automated Writing Evaluation) system that is the most well-known one among the AI tools that help people with learning languages. There have been several studies that show EFL learners’ perceptions of Grammarly as mostly positive and that they believe the program, besides pointing out mistakes in grammar, vocabulary, and mechanics, makes their writing more accurate overall (Dewi, 2022). The recent study of Dinçel and Saygı (2024), for example, considers more engagement factors and assesses how different proficiency levels and writing goals influence the use of Grammarly, thus indicating that students may have different interactions with the AI feedback according to their language and academic requirements. Simultaneously, surface-level correction, diminished attention to rhetorical organization, and the likelihood of students accepting suggestions uncritically, which might result in the loss of language awareness and critical thinking through the use of AI tools, have been some of the main concerns cited by educators and researchers when talking about the use of AI-assisted platforms in language learning.

The incorporation of AI in gamified platforms like Duolingo AI has brought up a whole new area in the academic discussion. According to Duolingo’s research and also by external studies, using adaptive algorithms along with the gamification features greatly enhances the teacher’s engagement as well as the students’ achievement in learning a language, especially in reading and vocabulary. Ouyang (2024), for instance, showed that there were apparent differences in student engagement (behavioral, affective, and cognitive) when using Duolingo as a supplementary tool (Ouyang, et.al., 2024). Still, most of the studies conducted so far rely on either large-scale learner data or quasi-experimental designs and serve to put effectiveness in the foreground over learners’ subtle perceptions of AI features within the app. The use of Duolingo’s AI-enhanced functions (e.g., smart tips, adaptive review, AI-generated exercises) by EFL university students has not received much qualitative attention so far.

There are several themes that cut across all these studies. The first one is that all the studies point to the same advantages of AI tools in the areas of self-direction, instant feedback, and personalized learning, which are in line with the learner-centered and constructivist pedagogies (Korkmaz, 2024). The second one is that the authors mention again and again issues like the accuracy and transparency of the AI feedback, the chance of plagiarism or academic dishonesty in case AI writes the text, and students getting along with AI and being less reluctant to face and solve linguistic problems (Alvarado & Andrade, 2025). The third one is that most of the literature looks at one particular tool or skill—writing,

vocabulary, or speaking—while AI-assisted learning is really a combination of practices that involves several platforms, devices, and contexts.

In addition, the authors note that the literature has a certain methodological bias consisting of quantitative survey studies and experimental interventions, while qualitative investigations capturing the students' lived experiences, contextual decision-making, and reflections on the ethical dimensions of AI use are rarer. There are some qualitative studies like those examining ChatGPT use in small samples or AI-enabled speaking applications—still the field lacks thick descriptions of how EFL learners navigate through the advantages and disadvantages of being able to use several AI tools in their everyday academic settings (Nouria & Fatima, 2024). More specifically, there are not many studies that investigate tools like ChatGPT, Grammarly, and Duolingo AI simultaneously in a single research project and explore the stance of the students regarding these technologies in terms of independence, critical thinking, and academic integrity.

The research presented here fills in the missing pieces by providing a qualitative study through interviews about the attitudes of EFL university students towards various AI-assisted language learning platforms. The research targets students who are regular users of ChatGPT, Grammarly, and Duolingo AI, hence it aims to paint a more complete picture of the perceived roles of the different AI tools as either supportive or opposing, and how the users express both the eagerness and the concerns regarding the pedagogical and ethical aspects. The authors argue that their research is not only revealing the learners' voices but also opening up the debates on the role of AI in language education in a manner that necessitates balanced and critically informed technological integration into EFL programs.

Method

1.1. Research Design

The qualitative research design was used in this study to examine the attitudes of English as Foreign Language (EFL) university students towards the adoption of Artificial Intelligence (AI) tools in language learning. A qualitative method was considered to be proper as it gives an opportunity to have an in-depth understanding of the participants' subjective experiences, beliefs, and concerns (Creswell & Poth, 2018). For conducting the research, semi-structured interviews were chosen as the research method, which allowed probing for more information from the participants without losing the consistency of the interviews. The adoption of this design corresponds to the recent calls for interpretive research that involves learners' comprehension of AI-assisted learning environments.

1.2. Participants

The research included 15 undergraduate students who were studying English as a foreign language in a public university's English language program. The selection of the participants was made with the purposive sampling technique to make sure that every single person was already familiar with AI-assisted language learning tools, such as ChatGPT, Grammarly, or Duolingo AI. The age of the students was between 18 and 23, and their English language proficiency levels were from intermediate (B1) to upper-intermediate (B2). Voluntary participation was allowed and no payment was given to the students. Before the data collection process began, all the participants had to fill out the informed consent form.

1.3. Data Collection Instruments

The key method for collecting data was semi-structured interviews. A questionnaire with open-ended questions was prepared with the help of existing literature on AI in language learning and student views. The guide addressed issues such as students' general feelings about AI tools, their use of particular platforms, and the advantages and challenges they saw, and it also brought up questions of ethics or academic standards in relation to AI use. Deeper insights were gained by adding follow-up questions whenever necessary.

1.4. Data Collection Procedure

Interviews, in total, were carried out during a three-week span in the spring semester. Each interview took roughly 30–45 minutes, and English was used as a language of communication to ensure consistency and to give a chance to the participants to express their experiences in the target language. Participants could choose whether to have the interviews conducted in a silent room on campus or through an online meeting service. All interviews were audio-recorded and later transcribed word for word with the consent of the participants. There were field notes created during and right after each interview to record conditions and first thoughts.

1.5. Data Analysis

Thematic analysis was the method used for data analysis in the study, following the process recommended by Braun and Clarke (2006). The analysis entailed six stages: (1) getting to know the data by means of reading the transcripts several times, (2) coding the data initially by pointing out the most important statements and/or the most prominent patterns, (3) grouping the codes together to come up with possible themes, (4) checking and fine-tuning the themes for their consistency and their connection to the research questions, (5) theme definition and naming, and (6) writing the final narrative. The coding was done manually. To support the legitimacy of the analysis, a second researcher was asked to review a part of the transcripts and the coding choices done independently.

Discussions were held regarding the differences in opinion until an agreement was reached. The university's research ethics committee gave ethical approval for the study. The researchers provided participants with research purposes, their rights that they could withdraw anytime, and confidentiality measures. Transcripts with personal identifiers were destroyed and pseudonyms were used instead. Audio files and transcripts were locked away in a safe place with a password that only the researcher knows. The study complied with the usual ethical principles for qualitative research in education.

With the aim of making the research more trustworthy, the study used techniques that fitted Lincoln and Guba's (1985) standards. Credibility was achieved through member checking where the participants went through the summaries of their interviews and pointed out where the recordings were wrong. The aspect of transferability was tackled by presenting in-depth descriptions of the research context and the attributes of the respondents. The dependability factor was taken care of through meticulous note-taking of the research process whereas confirmability was made more robust through the keeping of reflective notes and an audit trail during the analysis of the data.

Results

The thematic analysis of the semi-structured interviews produced four main themes corresponding to the participants' attitudes and experiences with AI-assisted language learning tools: (1) Perceived Advantages of AI Tools, (2) Enhancement of Learner Autonomy, (3) Concerns Regarding Overreliance and Critical Thinking, and (4) Ethical Considerations and Academic Integrity. These themes portray the intricate and at times contradictory nature of the students' interpretation of AI's contribution to their language learning process.

1. Perceived Advantages of AI Tools

A majority of the interviewees exhibited very positive attitudes towards AI tools and pointed out their roles in writing accuracy enhancement, instant feedback obtaining, and linguistic doubts clearing. Grammarly and ChatGPT among others were termed "supportive," "time-saving," and "reliable for quick help."

Out of the several students who remarked upon it, one felt that the AI feedback was quicker and more comprehensive than that of the traditional instructor. The following is what one of the interviewees said:

"When I compose an article, Grammarly works out my errors right away. I get to know my mistakes before they are erased from my memory." (Participant 4)

In a like manner, the learners appreciated ChatGPT's capability to elucidate the grammar rules or produce the instances:

"ChatGPT comes to my rescue when I am unable to think of instances. It makes things clear in a very basic manner." (Participant 9)

Duolingo AI won admiration for its interactive, gamified atmosphere and personalized review system. All in all, the participants pointed out a betterment in vocabulary, grammar, and writing fluency which they claimed to be the result of the speed and convenience of the AI tools.

2. Enhancement of Learner Autonomy

Another significant idea was the use of AI in promoting self-directed learning. The majority of the participants indicated that they were independent and self-assured in the use of AI tools, particularly in out-of-class situations. They recognized that AI was like having a language assistant "24/7" and that they could study according to their preferences.

A participant expressed his opinion in this way:

"Having AI means I don't have to be dependent on the teacher anymore. I can verify my writing at any moment. This situation forces me to take charge of my learning." (Participant 7)

Moreover, AI tools were considered to be an addition and reinforcement of teaching in the classroom. The students reported practicing, getting clarifications, or editing with AI after the lesson:

"After our lesson, I go through Duolingo and ChatGPT Every day. It helps me to retain what I have learned." (Participant 12)

Thus, AI was not seen as a substitute for teachers but rather as a part of the learning setting that helped develop independence and facilitated the process of learning all the time.

3. Concerns Regarding Overreliance and Critical Thinking

While the overall feedback was positive, a lot of the respondents still had worries about the level of their future dependence on AI. They thought that always relying on the computer's suggestions could possibly make their problem-solving, language and thinking skills weaker, and would reduce their human-like thinking capability.

One of the participants mentioned:

"I have a fear that if I rely too much on AI, I will not be able to produce a text on my own."
(Participant 3)

Another one talked about the AI-generated text really easily being accepted without any consideration:

"At times, it provides ready answers. It seems like I am cheating myself. I do not get the in-depth understanding when it is doing everything." (Participant 11)

Moreover, the participants were also doubtful regarding the AI's output quality, they said that AI does help to some extent, but it is "not always right", thus it is important for the students to always keep the critical attitude.

Their worries reflect a supposed conflict between the use of AI as a liberating force and the use of it as an impediment to cognitive development through self-reliance.

4. Ethical Considerations and Academic Integrity

A last topic was connected to the issue of ethics and academic honesty. Several users pointed out that AI could be misapplied and be used to write whole assignments, that would lead to the concerns of plagiarism and breaking of academic rules.

A user gave this example:

"Some students use ChatGPT to write everything. It's not fair, and it's not learning."
(Participant 14)

Others were in doubt about the use of AI in university policies:

"I don't know how much we are allowed to use AI. Sometimes I'm scared the teacher will think I copied." (Participant 2)

The users regarded AI as a good help for learning, but they insisted on the need for definite instructions and responsible use. Besides that, they wanted the learning environment where AI would be an addition—not a replacement—to human teaching and genuine students' effort.

In general, the results indicate mainly positive attitudes toward AI tools, whereby the students see their capability to boost writing quality, give instant feedback, and support learning through self-study. Nevertheless, the outcomes also bring up significant issues about dependence, lack of critical thinking, and ethical considerations. Such mixed feelings signify the necessity for proper AI integration and the demand for giving clear instructions on responsible and critical use of AI in academic context.

Discussion

The study focused on investigating EFL university students' attitudes towards the use of AI-assisted language learning tools and getting a clearer picture of the benefits, challenges, and ethical issues as seen by the learners. The results pointed to a large-scale positive inclination towards AI tools, mainly because of the writing assistance, instant feedback, and self-directed learning attributes. On the other hand, reliability on such tools, loss of critical thinking, and issues relating to academic integrity were some of the drawbacks that the respondents were worried about. The conclusions made in this regard really affirm and also add to the ongoing research in the area of AI application in language teaching and learning.

Perception and identification of AI tools' advantages by the participants of this study were in line with the previous studies that have pointed out the potential of automated feedback and intelligent tutoring systems to enhance teaching and learning. The participants' views on the instant and clear feedback from AI tools like Grammarly and ChatGPT are in line with the observations made by Wang and Vasquez (2020), who said that AI-supported writing environments make learners accurate in

language use because of the fast and personalized guidance provided. Likewise, positive attitude of the students toward ChatGPT as a source of explanations and examples is consistent with the findings of Xiao (2023), where learners regarded conversational AI as a good friend for brainstorming and language clearing. Participant feedback on Duolingo AI's customized functionalities was also found to be in line with the findings of the researchers who have described the role of gamified AI language solutions in increasing students' motivation and engagement (Ouyang, 2024).

Apart from the advancements in specific skills, the current dose of research points out the contribution of AI tools to learner autonomy. A number of students reported their feeling of being more accountable for their progress with the use of AI tools, a notion that is in line with Dinçel and Saygi's(2024) assertion that AI can be a facilitator of self-control over learning by permitting the learners to practice and revise independently. In this respect, AI is regarded as a supplement of the learning atmosphere, thus allowing the learners to be engaged continuously even outside the classroom. The aforementioned is a factor that strengthens the concept of AI as a trigger for the learner-centered teaching methodology, which in turn offers the flexible and personalized learning paths (Li, 2022).

On the other hand, notwithstanding the benefits, the study's respondents still voiced their concerns concerning the overuse of AI, lack of critical thinking, and the issue of getting wrong feedback. These fears mirror the discussion in the literature concerning the cognitive hazards related to over-reliance on automated support. Kohnke (2023) warns that the indiscriminate acceptance of AI-generated answers could lead to the development of weak analytical skills and hinder the ability of the learners to solve linguistic problems independently. The participants' remarks on "not knowing how to write by myself anymore" point to this conflict, thereby inferring that the learners know the possible disadvantages of entrusting too many cognitive functions to AI. This acknowledgment represents a subtle, contemplative comprehension of the ambivalent character of AI—its being a promoter of learning and, at the same time, a possible obstacle for learning at a deeper level.

Moreover, the ethical reservations expressed by the participants are in line with the ongoing debates about the integrity of academic practices in the learning environments supported by AI. The worry that students might use AI to completely write their papers is in line with the latest research that points out the educators' main concerns about the plagiarism and the extinction of the real students' efforts (El-Sahhar & Zaki, 2023). The participants' clarity about the institutional policies indicates the need for extensive clarification of AI usage as pointed out by Zawacki-Richter et al. (2019). In such case where no clear regulations or training in "AI literacy" are offered, the students might find it difficult to decide when they are legitimately, ethically, or pedagogically getting AI assistance.

One of the main benefits from this research comes from the comprehensive viewpoint it has taken on the students' experiences with the AI technology of ChatGPT, Grammarly, and Duolingo AI. While most of the earlier studies are limited to one tool or a special area of language, the present study makes a point that EFL (English as a Foreign Language) students interact with AI in various ways. Different students need different tools for different purposes—writing, the practice of language, the clarification of points—and this mixed-use results in a complex attitude that is a mixture of enthusiasm and caution. This finding broadens the literature by showing that the learners do not see AI as a single thing but as a whole ecosystem of tools that have different capacities and limitations.

The next big thing to do with AI was the balance that the students wanted to keep between the use of AI and the teaching of humans. Even though students appreciated AI, none of them considered it a substitute for teachers. On the contrary, the participants often pointed out the vital role of teachers in the face of human communication, personalized feedback, and moral counseling. This is in line with the opinion of Ouyang (2024) who insisted that AI should be seen as a supportive resource in mixed learning environments rather than a downfall of teaching mastery. Thus, findings bring forth the idea

that teachers have to use AI in such a way that the opportunities for critical thinking, dialogue, and human interaction are not lost.

The study's findings, in brief, present a complicated yet encouraging scenario of AI use in the EFL learning process. Even though the students admit a lot of input from AI in the pedagogic area, they, however, do not overlook its dangers, the main ones being dependency and academic dishonesty. These implications point to the fact that it is the dealing with risks and the adoption of technologies that will determine the success of AI application in language teaching and learning. Educators should direct students toward making use of AI in a critical and ethical way, and educational institutions ought to set up clear-cut policies to facilitate this irresponsible practice. The findings also bring to the foreground the necessity of AI literacy being an integral part of language teaching so that the students are able to judge the AI output, keep their independence, and not compromise their academic standards.

Conclusions

This research investigated EFL University students' attitudes towards AI-assisted language learning tools through the medium of semi-structured interviews, involving 15 participants. The results indicated that students, in general, considered AI tools like ChatGPT, Grammarly, and Duolingo AI as helpful aids that made their language learning easier especially by making writing accurate, giving feedback instantly and enabling self-study. The students' good views about AI confirmed the literature that recognized the potential of AI technology to bring about change in second language learning environments.

At the same time, students brought forth major concerns about relying too much on AI, loss of critical thinking, and the ethical issues raised by academic integrity concerns. These concerns indicate the necessity for critical engagement with AI tools rather than unreflective use. Therefore, the results accentuate the dual aspect of AI in language education: on the one hand, AI is capable of considerably supporting personalized and effective learning; on the other hand, its incorporation must be managed very carefully so that learners' cognitive participation is not impaired and academic standards are not lowered.

In summary, the research plays an important role in the ongoing discourse around the function of AI in EFL teaching by providing deep qualitative insights into the students' ways of coping with the pros and cons that come with different AI platforms. It points to a need for finding a delicate spot between the technical innovation and the pedagogical responsibility, making sure that the AI tools are the new supplements that add to the very basic elements of language learning, like human guidance, critical reflection, and communication that is real.

References

Alvarado, M. V. G., & Andrade, S. S. N. (2025). The impact of artificial intelligence on English language learning: A systematic review of tools, methods, and outcomes in language skills. *RUNAS. Journal of Education & Culture*, 6(12), 1.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

Chou, P. N., Chen, W. F., & Wu, H. C. (2023). An exploration of students' perceptions of AI-assisted language learning. *Education and Information Technologies*, 28(1), 1123–1141. <https://doi.org/10.1007/s10639-022-11407-z>

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Dewi, U. (2022). Grammarly as automated writing evaluation: Its effectiveness from EFL students' perceptions. *Lingua Cultura*, 16(2), 155–161. <https://doi.org/10.21512/lc.v16i2.8315>

Dinçel, K., & Saygı, C. (2024). The role of Grammarly® in EFL learners' writing development: A mixed-methods investigation. *Language Learning & Technology*, 28(2), 134–155. <https://doi.org/10.10125/123456>

El-Sahhar, H., & Zaki, R. (2023). University students' acceptance of AI writing tools: Perceptions, benefits, and challenges. *Computers & Education: Artificial Intelligence*, 4, 100138. <https://doi.org/10.1016/j.caeai.2023.100138>

Kohnke, L. (2023). Exploring the implications of AI tools for academic integrity in ESL contexts. *System*, 113, 102977. <https://doi.org/10.1016/j.system.2023.102977>

Korkmaz, H., & Akbiyik, M. (2024). Unlocking the potential: Attitudes of tertiary level EFL learners towards using AI in language learning. *Participatory Educational Research*, 11(6), 1–19. <http://dx.doi.org/10.17275/per.24.76.11.6>

Korkmaz, S. (2024). University students' attitudes toward artificial intelligence in foreign language learning. *Education and Information Technologies*, 29(1), 345–365. <https://doi.org/10.1007/s10639-023-11817-6>

Li, Z. (2022). Artificial intelligence in language education: A systematic review. *ReCALL*, 34(3), 282–301. <https://doi.org/10.1017/S0958344022000103>

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

Nouria, G. H. O. U. L., & Fatima, T. O. C. H. E. N. I. (2024). Exploring EFL students' attitudes towards AI-powered language learning applications for speaking skills development (Doctoral dissertation, University Center of Abdalhafid Boussouf-MILA).

Ouyang, F. (2024). The impact of Duolingo on EFL learners' behavioral, cognitive, and emotional engagement. *Computer Assisted Language Learning*, 37(2), 403–425. <https://doi.org/10.1080/09588221.2022.2122345>

Ouyang, Z., Jiang, Y., & Liu, H. (2024). The effects of Duolingo, an AI-integrated technology, on EFL learners' willingness to communicate and engagement in online classes. *The International Review of Research in Open and Distributed Learning*, 25(3), 97–115. <https://doi.org/10.19173/irrodl.v25i3.7677>

Rodriguez, L. B. Z. (2025). Chatbots and vocabulary learning: Perceptions from EFL university students. *UNESUM-Ciencias: Revista Científica Multidisciplinaria*, 9(3). <https://doi.org/10.47230/unesum-ciencias.v9.n3.2025.176-185>

Scherer, R., Siddiq, F., & Sánchez Viveros, B. (2019). The technology acceptance model in education: A systematic review. *Computers & Education*, 128, 13–35. <https://doi.org/10.1016/j.compedu.2018.09.017>

Wang, Y., & Vasquez, C. (2020). Automated writing evaluation in EFL writing: A meta-analysis. *Journal of Second Language Writing*, 47, 100713. <https://doi.org/10.1016/j.jslw.2019.100713>

Xiao, Y. (2023). University EFL students' perceptions of using ChatGPT for English learning: A qualitative study. *Computers and Education: Artificial Intelligence*, 5, 100150. <https://doi.org/10.1016/j.caeai.2023.100150>

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>

Zhi, X. (2024). Students' attitudes toward artificial intelligence in EFL classrooms: The role of teacher-student relationships. *System*, 120, 103125. <https://doi.org/10.1016/j.system.2024.103125>

AUTHOR BIODATA

Ziyoda Khalmatova is a PhD, senior lecturer of the Foreign Languages Teaching Department at Khoja Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan. She gained her PhD degree in Institute of Turkic Studies at Hacettepe University, Türkiye. She is interested in teaching and learning strategies, models and methods using new technology. ORCID: <https://orcid.org/0000-0001-7695-5282> Email: ziyoda.khalmatova@ayu.edu.kz.

The ways in which AI is changing our ELT practices

Prof. Dr. Abdülvahit ÇAKIR

Ufuk University, Ankara, Türkiye

APA Citation:

Çakır, A. (2025). The ways in which AI is changing our ELT practices

Abstract

The integration of artificial intelligence (AI), particularly generative AI tools like ChatGPT, into English Language Teaching (ELT) is proceeding at an unprecedented rate. While these technologies offer transformative potential, their rapid adoption often outpaces critical evaluation of their practical applications and inherent limitations. This paper critically examines prevalent assumptions in the ELT community, including the notions that AI possesses superior intelligence, can replace human educators, will soon be free of factual inaccuracies ("hallucinations"), or will invariably level the global educational playing field. Through concrete examples—such as ChatGPT's capacity to generate confident yet fabricated explanations—the article demonstrates how these beliefs can be misleading and pedagogically counterproductive. In response, the study underscores the critical importance of a "human-in-the-loop" approach, concluding with practical recommendations for the ethical integration of AI in ELT, outlining strategic imperatives for pedagogy and assessment while highlighting potential pitfalls to avoid.

Keywords: Artificial Intelligence (AI), ChatGPT, English Language Teaching (ELT), Hallucination, Digital Divide, Human-in-the-Loop

We should take informed decisions about AI in ELT.

AI tools like ChatGPT have become more popular in the past few years than almost any other new technology. People from many areas of life, including education, have started using them very quickly. By May 2024, research showed that 82% of university students and 72% of schoolchildren in the United States had used AI for their studies, and many said they used it several times a week. The situation is quite similar in other developed and developing countries, including Turkey.

Research by Hui Rong and Charlene Chun (2024) shows that university students (from bachelor, masters, and doctorate levels) use AI primarily for searching for information (69%), followed by checking grammar (42%), summarizing documents (33%), paraphrasing documents (28%), and creating first drafts (24%). With 66% of students using it in their studies, ChatGPT stands out as the most commonly used tool. Grammarly and Microsoft Copilot also show notable levels of adoption, each used by 25% of students. According to the same survey the top concerns among students about a university's use of AI are privacy and data security (61%), trustworthiness of AI generated content (51%), bias and fairness in AI evaluations and decisions (32%) and insufficient AI capabilities amongst faculty (28%).

Should we be worried? Some people think so. Many teachers are afraid that students might use AI to cheat instead of learning properly, or that they might trust false or unreliable information produced by these tools. Others, however, are more positive. They believe AI can make learning easier and more personal, while also helping teachers by saving time and giving new teaching opportunities.

What is certain is that education will change. Teachers, students, parents, school leaders, and policymakers all need to respond to these changes in order to get the best results. The real question is not only what AI can do, but how we choose to use it for better learning.

According to Prof. Bharat N. Anand (2025) at Harvard University, the rise of technologies and the time it took for different technologies to reach 50% penetration in the U.S. economy.

- Computers: It took 20 years to reach about 30% penetration.
- Radio: It took about 20 years to reach half the population.
- TV: 12 years
- Smart phones: 7 years
- Smart speakers like Siri and Google assistant: 7 years
- Chat bots: 2,5 years

The above durations clearly show that new digital technology tools are apparently becoming popular faster and faster.

AI is not something new as some of us might think. It has been around for more than half a century in different forms.

- Machine learning (Probability theory, etc. before 1950s)
- Deep learning (Key breakthrough in 1986)
- Transformers (Key breakthrough in 2017)

As laymen, ordinary people, we have started to develop some shared views about digital tools like ChatGPT and DeepSeek. Let's have a look at some of them.

1. Generative AI's transformative potential stems from its "intelligence capacity" viz., how smart its output is.
2. Prudent educators should wait until the generative AI output – its intelligence gets better and it's less prone to wrong answers.
3. Bot tutors are unlikely to be superior to well-honed active learning methods in the classroom.
4. Generative AI will ultimately level the playing field in education.

The First View

There has been gradual improvement over time. There were some discrete improvements in intelligence two years ago as opposed to the improvement in access or the interface that we have with the intelligence. What do we mean by that?

We can divide the history of human communication roughly into four stages:

1. Sitting around camp fires talking to each other
2. Writing pictures on the walls, that was graphics
3. Writing scrolls (papyrus, leather or parchment) i.e., texts, books etc.
4. Using ones and zeros – mathematics

Basically this was the order of developments in human communication, first there used to be only **talking**. The primary channel of communication was speech. Later **pictures** / graphics appeared in addition to speech. Graphics did not die on the wind, it kept messages for longer periods of time. And

then in the third stage, **written texts** emerged marking the beginning of history. Last the use of **ones and zeros** (mathematics) in software became dominant.

Developments in communication using digital technology have reversed this historical order. 70 years ago we were using punch cards i.e., ones and zeros, which made the representation of all numerical values and logical information possible (Stage 4). Then we moved on to DOS prompt commands, which were nothing other than written words like “print” (stage 3). Then with the advent of interfaces we began using symbols, graphics and icons, which enabled ordinary people and even children to use computers effectively (stage 2). And finally with AI tools like ChatGPT we began giving oral prompts, simply by talking just the same way as we did before the advent of pictures and writing (stage 1).

What is actually happening is that the distance between computers and human beings is shrinking. Early computers required complex commands and programming languages. Now chatbots, voice assistants and AI help us talk to a machine. We just say “Hey Siri, remind me to call my daughter at 9 tomorrow morning”, and it does. So AI is miraculous not because of its intelligence or the precise information it produces but because of the speed and ease with which humans reach information through it.

The Second View

Generative AI does not always produce correct information. Currently some content produced by ChatGPT, for example, is factually incorrect, fabricated, or misleading. Such cases are called “hallucinations”. For example, generative AI may invent citations by giving a journal article or a web address that does not exist or add non-existent characters to a historical event. Some people believe that “AI often produce sub-par results — false citations, inaccurate statements” (Kwon 2025: 576). Therefore, many people naturally look at such hallucinations and say let’s wait, let’s wait till it gets better. However, hallucinations are a fundamental intrinsic property of generative AI because they are probabilistic models, meaning that there may not come a day when we have AI applications which produce no false information. What we can do is use GenAI responsibly. In digital education teachers and students must understand the risks of the genAI and the need for human-in-the-loop systems instead of waiting for hallucination free AI models. This approach enables the integration of emerging technologies while safeguarding educational standards (Ciubotaru, 2025).

Think of a big company receiving tens of thousands of calls every day. They are answered automatically with AI. Even if 3 % of the automated answers are wrong, the cost is acceptable, the company will lose not much. But in areas like medicine or law, even 1% errors may cost too much. A patient may lose his or her life because of a tiny percentage of errors made by AI in the medication prescribed. Similarly, in a legal trial 1% of errors made in one’s hearing may result in losing the case altogether. It is a matter of trade-off between cost and benefits. What are the benefits of using generative AI and what is the cost of its errors? Cost of errors, as the examples above show, changes from domain to domain. In some cases, it is high. In some others it is low.

The Third View

We may tend to believe that bot tutors are unlikely to be as effective as well-honed active learning methods in the classroom. Simply put, bot tutors cannot be as good as human instructors. Interestingly the scores of the students using the AI bots were higher than the human tutors although they are no ordinary tutors and they have been refining their craft year in and year out. What was even more surprising, the engagement was higher in bot tutors’ classes. What can we conclude from this experiment? Students probably felt much freer with bot tutors and asked whatever they did not know without the risk of being embarrassed resulting in better learning.

The Fourth View

Looking at the amazing developments we may naturally think that GenAI will level the playing field in education. Everybody even people in remote villages have access to the same kind of technologies as those in elite universities. And this is going to level everything. There is a possibility it might go exactly the other way. The benefits might prove disproportionately to those who have domain expertise and a knowledge of a subject. You start using your judgement. In some domains, the gap between high performance high knowledge workers and no knowledge workers is increasing. You are basically using your judgement to filter out what is useful and what is not useful. If I didn't know anything about the subject, I basically don't know what I don't know. Prompts are about garbage in garbage out.

So language teachers who use AI tools are well advised to not to ask “what is AI going to do to me?” but rather ask “which are the tasks I can automate and which are the tasks that I don't want to touch?”.

What does GenAI know?

I asked ChatGPT the following question. In Turkish why do we say "tahliye borusu", "sınav sorusu", "çocuk korosu" but "elma suyu" not "elma susu"?

ChatGPT's Answer:

Great question. This goes right to the heart of **Turkish noun compounds (isim tamlamaları)** and why sometimes we see **-sı/-si/-su/-sü** and sometimes we don't.

1. Two main types of compounds in Turkish

a) Belirtisiz isim tamlaması (non-possessive compound)

- Structure: **[Noun + Noun(+sI)]**
- The second word gets the **3rd person singular possessive suffix**.
- Examples:

tahliye borusu → “drain pipe” (pipe for drainage)

sınav sorusu → “exam question” (question of the exam)

çocuk korosu → “children's choir” (choir of children)

Meaning: The first noun defines/limits the second, and the possessive suffix shows this relationship.

b) Basit birleşik isim (compound noun without suffix)

- Structure: **[Noun + Noun]** (no -sI on the second noun).
- Examples:

elma suyu → “apple juice”

portakal kabuğu → “orange peel”

güneş ışığı → “sunlight”

Here, the first noun directly describes the material, content, or source of the second noun. The possessive suffix is **not used** because it sounds redundant or unnatural.

2. Why not *elma susu*?

Because *su* behaves differently from nouns like *boru*, *soru*, *koro*:

- *su* is seen as a **substance/content word** rather than an object/container.
- When you say *elma suyu*, you mean “the juice **made from apple**.”
- If you said *elma susu*, it would sound like “the water **that belongs to the apple**” → unnatural in Turkish. What about, then, “elma posası” or “elma lekesi”?

So:

- *tahliye borusu* = “the pipe of drainage” (container/object relation).
- *sinav sorusu* = “the question of the exam.”
- *çocuk korosu* = “the choir of children.”
- *elma suyu* = “the juice from apple” (material/content, not possession).

Rule of Thumb

- If the relationship is **ownership, belonging, or functional**, use **-sI** → *sinav sorusu*, *çocuk odası*, *araba lastiği*.
- If the relationship is **material, content, source, or kind of substance**, use **no -sI** → *elma suyu*, *altın yüzük*, *pamuk ipliği*

The above explanation by ChatGPT of how a noun modifies another noun in Turkish, erroneous all through except for the English equivalents of the noun phrases, constitutes a perfect example of hallucination.

- **He who knows not, and knows not that he knows not, is a fool; shun him.**
- **He who knows not, and knows that he knows not, is a student; teach him.**
- **He who knows, and knows not that he knows, is asleep; wake him.**
- **He who knows, and knows that he knows, is wise; follow him.**

Which of these characters, do you think, ChatGPT is like? Maybe the second?

As human beings, we may be tempted to liken Generative AI to one of the types of people mentioned above. However, such an analogy would be baseless and misleading. GenAI does not possess knowledge or awareness; it merely predicts the most probable sequence of words by performing statistical analyses on the texts available to it in relation to the prompt. Yet, it speaks with remarkable confidence, presenting its answers assertively and giving the impression that it truly understands the topic. This tone is not accidental—it reflects a design choice by its creators, who prioritized clarity and authority over hesitation.

DIGITAL DIVIDE vs FLAT WORLD

The concept of flat world

Thomas Friedman uses the metaphor of a “flat world” to describe globalization and the equalizing effect of modern technology. In *The World Is Flat* (2005), he argues that the combined forces of

globalization and technological advancement have largely levelled the playing field. According to Friedman, the digital revolution enables people across the globe to access reliable information with unprecedented speed, reducing traditional barriers and helping to minimize educational disadvantage and inequality.

The concept of digital divide

The term *digital divide* first appeared in the reports published by U.S. Department of Commerce in 1995-96. The digital divide theory claims that there is an inequality between people in terms of their access to digital technologies, their ability to use them, and their domain knowledge. This inequality widens the social and economic gap between them, which is openly incompatible with what the optimistic flat world theory says.

Some people have access to technologies like the internet, computers and smartphones, but some do not for one reason or another. So there is certainly an inequality between these groups. We also observe that some of the people who have access to such technologies are not effective users either because they lack advanced computer skills or because their knowledge about the domain of research is limited.

The edX platform, a Harvard and MIT joint creation, is a good example. It was launched in 2012. It offered high quality courses online. The purpose was to reduce the digital divide in education. They have more than 3000 courses now (MOOCs = Massive Online Open Courses). They are all free if no certificate is demanded. Millions of learners take these courses. But only less than 5% complete the courses they take. 75 % of those who successfully complete the courses were already university graduates. The better educated learners are getting even better in these programs, deepening the digital divide.

Similarly, Generative AI (GenAI) could **widen the gap** between people who are equipped with strong digital skills, knowledge, or expertise and those who are not. The benefits of GenAI (such as increased productivity, better problem-solving, or economic advantages) are more likely to go to people who **already know how to use technology well or have specialized knowledge in their field**. So instead of helping everyone equally, GenAI may make experts even more powerful while leaving behind those without the same background, thus **worsening the digital divide**.

This is valid in the field of English language teaching, too. It goes without saying that learners with a better command of English and familiarity with research methods and digital tools will profit from AI tools more than those without. *Seterus paribus*, of two learners, the one equipped with necessary knowledge and skills will gain more from a GenAI like ChatGPT leaving the other behind. Those who can formulate good prompts will get the desired information from AI tools and the formulation of good prompts takes knowledge of the area and digital skills. Garbage in garbage out. Knowledge of the field is crucial not only for good prompts but also for judging the answers one gets. Otherwise users of AI might get and use inappropriate outcomes.

What are the implications of such a divide? What should instructors do about it?

- Generative AI's transformative potential stems from its ease of access viz., how accessible its input is.
- You can benefit from generative AI even when its intelligence is worse –because of the potential for cost and time-savings in low-cost-of-error environments.
- Bot tutors are already yielding effective results on learning and mastery.
- GenAI may increase the digital divide – its gains may accrue disproportionately to those who already have domain expertise.
- We need to layer technology onto existing workflows.

- What we urgently need is a strategic conversation about our role and purpose as teachers and educators.

What should a foreign language instructor do or should not do to narrow the digital divide between her students? What strategies and attitudes could she adopt?

Here are some **strategies and attitudes** a foreign language instructor could adopt to help narrow the digital divide between her students:

- To bridge the digital divide in the classroom, educators can adopt several targeted strategies. An effective initial step is to identify students with underdeveloped digital skills and implement peer-assisted learning through structured group work. This approach allows more proficient students to mentor their peers, fostering a collaborative learning environment.
- Furthermore, teachers should proactively introduce students to essential digital tools, such as writing platforms, educational applications, and online dictionaries. When selecting these resources, it is crucial to prioritize those that are free and accessible to all students to ensure equity.
- In-class demonstrations are key to building competency. Educators should model how to use technology for practical tasks like checking grammar, verifying spelling, and enriching vocabulary with digital thesauruses. Finally, it is vital to create inclusive activities that incorporate artificial intelligence, ensuring every student—especially those who are less confident—actively participates. This not only builds their technical skills but also cultivates a genuine sense of accomplishment and digital self-efficacy.

In a nut shell: A foreign language teacher should do everything to close the gap between her students in terms of digital skills. If she supports those coming from behind and avoid intimidation, the digital divide will shrink. It is important that all students feel confident and be motivated to use digital tools for language learning.

Nowadays language students are inclined to ask ChatGPT to do any homework assigned to them. To what extent is this acceptable? Is this some kind of cheating? What could a prudent teacher do about it? I asked ChatGPT.

ChatGPT:

This is one of the key challenges in today's classrooms. Let me break it down clearly:

To what extent is this acceptable?

- **Acceptable use:** If students use ChatGPT as a *support tool* — for brainstorming, checking grammar, generating examples, or clarifying doubts — it can be a valuable aid to learning.
- **Problematic use:** If students simply copy-paste answers from ChatGPT and present them as their own work, they are not practicing the language skills themselves. This is closer to **cheating**, because they are avoiding the learning process.

So the acceptability depends on **how** they use the tool:

- *Supportive use = acceptable*

- *Substitutive use (doing all the work) = not acceptable*

Is it cheating?

- **Yes, if:** The assignment is meant to test their own writing, speaking, or comprehension, and they pass off AI output as their own.
- **No, if:** They are transparent about using it, and the task allows or encourages AI as a learning partner.

Cheating isn't only about rules; it's about **dishonesty** and **loss of learning opportunity**.

As prudent teachers, we can advise our students to:

- Be critical thinkers and verify any information they receive from ChatGPT.
- Use ChatGPT as a learning aid, not as a substitute for doing their own homework.
- Uphold academic honesty and never present texts generated by ChatGPT as their own work.
- Use ChatGPT to refine and develop their own ideas, rather than copying or merely paraphrasing its responses.
- Ask ChatGPT to simplify complex texts so that they can better understand materials beyond their current level.
- Express ideas in their own words, demonstrating genuine understanding.
- Be aware that ChatGPT can make mistakes and remain alert to possible inaccuracies or hallucinations.
- Remember that ChatGPT may not have access to the most recent information or developments.
- Keep in mind that ChatGPT lacks human experience and emotional understanding.
- value their own learning and creativity above all, recognizing that authentic effort outweighs merely repeating ChatGPT's words, however impressive they may sound.

This response from ChatGPT is quite reasonable. Accordingly, we can conclude that using ChatGPT—or any other AI tool—is **not necessarily a form of cheating**. Language teachers should allow the use of AI as long as learners clearly indicate which parts of their work were produced **with AI assistance and which were created independently**. Rather than banning AI outright, teachers should **explain and demonstrate how to use it effectively** to support the development of all four language skills.

AI Use Policy for Language Learning

Although Artificial Intelligence (AI) tools such as ChatGPT can greatly support language learning, there are still no clearly defined rules regarding their ethical use in completing homework or producing academic texts. In general, we can say that AI should be **encouraged** when it is used to **support and enhance learners' knowledge and skills**, but it should be **discouraged** when it replaces learners' own thinking and creativity, thereby limiting the development of higher-order thinking skills.

Based on recent experiences shared by practitioners, we as language teachers can formulate our own **AI use policy** along the following lines:

- **Encourage supportive use:** Allow learners to use AI tools to check the accuracy of their texts in terms of grammar, vocabulary, and spelling. Students can also use AI for brainstorming, generating new ideas, or practicing language skills through conversation.

- **Emphasize originality:** Make it clear that you value students' own work and answers, even if they are not as polished as AI-generated ones. Submitting AI-produced texts as one's own constitutes plagiarism, just like copying from traditional sources.
- **Warn about reliability issues:** Remind students that AI tools can produce inaccurate or fabricated information ("hallucinations"), especially when generating creative or original ideas. Encourage them to verify AI outputs rather than accept them blindly.
- **Promote genuine learning:** Explain to students that real progress comes from thinking, practicing, and making an effort themselves—not from letting AI or others do the work on their behalf.

References

Anand, A., & Bharat, N. (2025, March). *AI in education: Shaping the future of classrooms* [Video]. #IndiaTodayConclave2025. YouTube. <https://www.youtube.com/watch?v=ssZZJ5ArWLo>

Ciubotaru, B.-I. (2025). The hallucination problem in generative artificial intelligence: Accuracy and trust in digital learning. *International Conference on Virtual Learning*, 20, 35–45. <https://doi.org/10.58503/icvl-v20y20253>

Kwon, D. (2025). Scientists split on ethics of AI use. *Nature*, 641, 574–578. <https://www.nature.com/articles/d41586-025-01463-8>

Lee, C., et al. (2025). Generative AI risks and resilience: How users adapt to hallucination and privacy challenges. *Telematics and Informatics Reports*, 100221. <https://doi.org/10.1016/j.teler.2025.100221>

Moorhouse, B. L. (2024). Beginning and first-year language teachers' readiness for the generative AI age. *Computers and Education: Artificial Intelligence*, 6, Article 100201. <https://doi.org/10.1016/j.caeari.2024.100201>

Rong, H., & Chun, C. (2024). *Digital Education Council global AI student survey 2024: AI or not AI—What students want*. Digital Education Council.

Samuel, A., & Maciuba, J. M. (2025). Guidelines for the responsible use of generative AI in manuscripts submitted to JCEHP. *Journal of Continuing Education in the Health Professions*, 45(3), 141–144. <https://doi.org/10.1097/CEH.0000000000000612>

A Review of Willingness to Communicate (WTC) in L2 and Learner Psychology

Ahmet Selçuk Akdemir^a

^a *Malatya Turgut Özal University, Malatya, Turkey*

APA Citation:

Akdemir, A. S. (2025). A Review of Willingness to Communicate (WTC) in L2 and Learner Psychology.

Abstract

Willingness to Communicate (WTC) is defined, in short, as one's desire to start or involve in communication in a specific context. As WTC has made its way to L2 studies, several personal and external factors have been identified as effecting or being affected with WTC. Nowadays the psychology of language learner has gained popularity regarding its importance on language learning process. Happiness, mindset, language enjoyment, and positive atmosphere of language learning are among the main concepts taken into account regarding the psychology of language learner. Current paper is a review on WTC and its novel position in context with learner psychology. The study aims at re-positioning the role of WTC in learner psychology. To this end, current research trends and tendencies are reviewed. The paper concludes with practical and theoretical implications.

Keywords: WTC, learner psychology, applied linguistics.

1. Introduction: From Competence to Volition

For decades, Second Language Acquisition (SLA) research focused on communicative competence—what a learner can do. However, researchers observed a paradox: highly competent learners often remained silent, while less competent ones spoke freely. This led to the conceptualization of Willingness to Communicate (WTC). WTC is defined in L2 contexts as "a readiness to enter into discourse at a particular time with a specific person or persons, using a L2" (MacIntyre et al., 1998). It is currently considered the primary psychological gateway to L2 use; without WTC, communicative competence remains latent and unpractised. During last three decades WTC has evolved from its initial form. Current WTC is a dynamic concept being shaped by several personal and psychological factors.

2. The Heuristic Model (The Pyramid)

The foundational framework for L2 WTC is the Heuristic Model proposed by MacIntyre, Clément, Dörnyei, and Noels (1998). This model is visualized as a pyramid, illustrating that WTC is the final step before actual communication behavior.

The Layered Architecture:

Layers I & II (Proximal/Situational): These are the immediate drivers. The decision to speak depends on the state of the learner at that exact moment (e.g., "Do I feel confident talking to this person on this topic?").

Layers III & IV (Distal/Enduring): These include motivation (e.g., L2 Motivational Self System) and generalized self-confidence.

Layers V & VI (Foundational): Deeply rooted factors like personality (introversion/extraversion) and intergroup climate.

Key Insight: The model argues that psychological variables (anxiety, confidence) filter through situational variables to produce WTC.

3. Psychological Drivers of WTC

Research has identified a "Big Three" of psychological predictors for WTC: Anxiety, Self-Confidence, and Motivation.

A. Anxiety (The "Brake")

Foreign Language Classroom Anxiety (FLCA) is the most consistent negative predictor of WTC.

Mechanism: Anxiety acts as an affective filter. High anxiety consumes working memory and creates an "avoidance orientation," causing learners to remain silent even when they know the answer.

The "Anxiety-WTC" Loop: MacIntyre & Legatto (2011) found that anxiety is not just a cause but a result. A failed communication attempt lowers WTC for the next turn, creating a negative feedback loop.

B. Self-Confidence (The "Accelerator")

Often cited as the strongest predictor of WTC, L2 self-confidence is comprised of two sub-factors:

Perceived Competence: "I believe I can communicate effectively."

Lack of Anxiety: "I am comfortable communicating."

Literature consensus: High self-confidence can override low actual competence. A confident beginner typically has higher WTC than an anxious advanced learner.

C. Motivation (The "Fuel")

Modern research links WTC to Dörnyei's L2 Motivational Self System.

Ideal L2 Self: Learners who visualize themselves as successful L2 speakers have higher WTC because every interaction bridges the gap between their "current self" and "ideal self."

Integrativeness: A desire to integrate with the L2 community predicts higher WTC outside the classroom.

4. The "Dynamic Turn": WTC as a Complex System

Early research treated WTC as a stable trait (like personality). However, Dynamic Systems Theory (DST) has revolutionized the field, arguing that WTC fluctuates second-by-second.

The Idiodynamic Method

MacIntyre and Legatto (2011) pioneered the Idiodynamic Method, where learners watch videos of their own interactions and rate their WTC moment-by-moment.

Key Findings:

Volatility: WTC is not a flat line. It spikes and crashes within a single conversation based on variables like word retrieval difficulties, a smile from the interlocutor, or a sudden pause.

The Butterfly Effect: A minor negative cue (e.g., a teacher's frown) can cause a "phase shift," dropping WTC from high to zero instantly.

5. The "Positive Turn": Enjoyment and Grit

Since 2015, the rise of Positive Psychology (PP) in SLA has introduced new variables beyond simple "low anxiety."

Foreign Language Enjoyment (FLE)

Dewaele et al. (2019) and subsequent studies (e.g., Chen et al., 2025; Lee & Taylo, 2024; Yin & Zhou, 2025) have established FLE as a distinct positive predictor of WTC.

FLE vs. Anxiety: They are not opposites on a single spectrum. A learner can experience both high anxiety and high enjoyment (e.g., the thrill of a difficult challenge).

Impact: Enjoyment broadens the learner's "thought-action repertoire," making them more willing to take risks (WTC) despite potential errors.

L2 Grit

Recent studies (2020–2024) have examined Grit (passion and perseverance for long-term goals). Gritty learners maintain WTC even after repeated communicative failures, viewing them as necessary steps rather than embarrassments. Based on current research it is possible to suggest a flow as follows:

L2 Grit → Increased FLE → Higher WTC

In this sense, gritty learners often possess a clearer visualization of their "Ideal L2 Self" (the person they want to become). This vision sustains the motivation required to initiate communication (Derakhshan et al., 2022; Lan et al., 2021)

Table 1. Psychological Correlates of L2 WTC

Variable	Correlation with WTC	Mechanism
Self-Confidence	Strong Positive (++)	Reduces fear of negative evaluation.
Anxiety (FLCA)	Strong Negative (--)	Induces avoidance behavior.
Motivation (Ideal Self)	Positive (+)	Increases value of interaction.
L2 Grit	Positive (+)	Sustains WTC despite failure.
Enjoyment (FLE)	Positive (+)	Broadens attention; encourages risk-taking.

6. Conclusion

The scientific consensus on Willingness to Communicate has shifted from a static "personality trait" to a dynamic, situation-dependent psychological state (Cheng et al., 2025; Kirkpatrick et al., 2025; Lu, 2025). While linguistic competence is the engine of communication, psychology (anxiety, confidence, enjoyment) is the ignition. Current Academic Verdict: To increase L2 communication, educators should focus less on error correction (which may spike anxiety) and more on "Loving Pedagogy" and perceived competence, fostering an environment where WTC can survive the inevitable mistakes of learning.

References

Chen, X., Alruwaili, A. R., Azari Noughabi, M., Ghasemi, A., & Zhen, C. (2025). The mediating role of psychological capital in the relationship between EFL learners' L2 grit and L2 WTC. *Frontiers in Psychology*, 16, 1621340.

Cheng, Q., Sun, Z., & Chen, X. (2025). Chinese learners' willingness to communicate and foreign language anxiety in translanguaging practices: the idiodynamic perspective. *Journal of Multilingual and Multicultural Development*, 1-22.

Derakhshan, A., Fathi, J., & Nourzadeh, S. (2022). Testing a model of WTC based on Ideal L2 Self and grit among Learners of Persian as a Foreign Language. *Journal of Teaching Persian to Speakers of Other Languages*, 11(1), 29-50.

Dewaele, J. M., Chen, X., Padilla, A. M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in psychology*, 10, 2128.

Kirkpatrick, R., Derakhshan, A., Wang, Y., Taghizadeh, M. S., & Almuhan, M. A. (2025). Demystifying the role of L2 speaking embarrassment and anxiety in the interaction of bilingual learners' self-efficacy and willingness to communicate (WTC): A multinational serial mediation study. *Applied Linguistics Review*, (0).

Lan, G., Nikitina, L., & Woo, W. S. (2021). Ideal L2 self and willingness to communicate: A moderated mediation model of shyness and grit. *System*, 99, 102503.

Lee, J. S., & Taylor, T. (2024). Positive psychology constructs and extramural English as predictors of primary school students' willingness to communicate. *Journal of Multilingual and Multicultural Development*, 45(7), 2898-2916.

Lu, M. (2025). Exploring the relationships among emotions, willingness to communicate and flow experience in Chinese EFL learners: A structural equation modelling approach. *Journal of Multilingual and Multicultural Development*, 46(9), 2857-2872.

MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied linguistics*, 32(2), 149-171.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The modern language journal*, 82(4), 545-562.

Yin, K., & Zhou, L. (2025). The relative importance of peace of mind, grit, and classroom environment in predicting willingness to communicate among learners in multi-ethnic regions: a latent dominance analysis. *BMC psychology*, 13(1), 1-17.

AUTHOR BIODATA

Ahmet Selçuk Akdemir is an associate professor of English Language Teaching (ELT) at Malatya Turgut Özal University, Turkey. Dr. Akdemir's research interests cover L2 listening, WTC, ELT pedagogy and language teacher training. As well as conducting editorial roles in several academic journals, Dr. Akdemir teaches EFL courses at undergraduate and graduate levels.

Cultural Elements of 'Uptown': A Coursebook Analysis

Ahmet Selçuk Akdemir^a

^a *Malatya Turgut Özal University, Malatya, Turkey*

APA Citation:

Akdemir, A. S. (2025). Cultural Elements of 'Uptown': A Coursebook Analysis

Abstract:

In 2023, Ministry of National Education (MoNE) has announced the new education model to be implemented in Turkey: Türkiye Century Education Model (TCEM). This model is presented as a major change in educational policy and programmes for all school levels and courses. As MoNE states, this model is the outcome of an intensive preparation period since 2014. Several experts, academicians and teacher trainers have taken role in preparing the programmes, the coursebooks and educational documents for TCEM. Culture, values and moral issues are the basic highlights of TCEM. This paper aims at investigating the cultural elements of English coursebook 'Uptown', which is prepared for TCEM.

Keywords: TCEM, coursebook, culture.

1. Introduction

The landscape of education in Turkey has undergone a significant transformation with the introduction of the Türkiye Century Education Model (TCEM) (Türkiye Yüzyılı Maarif Modeli). Announced in 2023 by the Ministry of National Education (MoNE), this model represents a paradigm shift from previous curricula, aiming to raise individuals who are not only academically competent but also morally grounded and culturally aware.

The TCEM is described not merely as a curriculum update but as a holistic educational philosophy developed through an intensive preparation period beginning in 2014. It emphasizes the integration of "root values" (kök değerler)—such as justice, friendship, honesty, patience, and patriotism—into every aspect of instruction.

In the context of English Language Teaching (ELT), textbooks serve as the primary vehicle for curriculum delivery. They are not just linguistic manuals but cultural artifacts that convey values and social norms. The introduction of the new 9th-grade coursebook, 'Uptown', marks the first implementation of TCEM principles in high school ELT. This paper aims to analyze 'Uptown' to understand how the specific cultural and moral objectives of the TCEM are materialized in language learning materials.

2. Theoretical Framework

2.1. Culture in Language Teaching

Language and culture are inextricably linked. As noted by scholars like Byram (1997) and Kramsch (1993), it is impossible to teach a language without teaching culture. In the context of ELT, cultural content is often categorized into three types (Cortazzi & Jin, 1999):

Source Culture: The learner's own culture (e.g., Turkish culture).

Target Culture: The culture of English-speaking countries (e.g., UK, USA).

International Target Culture: Various cultures from around the world using English as a lingua franca.

2.2. *The Shift in TCEM*

Historically, ELT materials often leaned heavily on Target Culture. However, the TCEM explicitly highlights "culture, values, and moral issues," signaling a shift towards validating the Source Culture and using English to express national identity and universal moral values. This aligns with the concept of "Glocalization" in education—thinking globally while retaining local cultural integrity.

Glocalization—a portmanteau of globalization and localization—originally popularized in sociology by Roland Robertson (1995), refers to the adaptation of global ideas and practices to suit specific local contexts. In English Language Teaching (ELT), this concept has emerged as a critical response to the hegemony of "Western" or "Center" methodologies, advocating instead for pedagogical practices that respect local culture, linguistic reality, and learner needs.

The Shift from Method to Post-Method

The primary driver of glocalization in ELT is the critique of the "one-size-fits-all" approach. Historically, methods developed in the Anglophone West, such as Communicative Language Teaching (CLT), were exported globally with little regard for local constraints.

The Post-Method Condition: Kumaravadivelu (1994, 2006) is a seminal figure in this discourse. He argues against the search for a perfect "global method," proposing instead a "post-method pedagogy" centered on particularity. This requires teachers to be sensitive to a particular group of teachers teaching a particular group of learners pursuing a particular set of goals within a particular institutional context.

Reclaiming the Local: Canagarajah (2005) emphasizes "reclaiming the local" in language policy and practice. He argues that local knowledge and vernacular practices (such as code-switching) should not be viewed as deficits but as legitimate assets in the classroom.

Materials and Curriculum: The "Global" Textbook

A significant area of tension in glocalization is the use of international coursebooks produced by major Western publishers.

Cultural Inappropriateness: Gray (2002) critiques global coursebooks for promoting a "grammar of culture" that represents a sleek, aspirational, and often exclusively Western lifestyle, which may alienate learners in diverse contexts.

Glocalizing Materials: To counter this, scholars like Tomlinson (2011) and Shin (2006) advocate for context-sensitive materials. This involves "glocalizing" textbooks—keeping the global linguistic target (English) but replacing Western cultural content with local characters, settings, and cultural norms to increase learner engagement and identity investment.

English as an International Language (EIL)

Glocalization is intrinsically linked to the conceptual shift from English as a foreign language to English as an International Language (EIL) or World Englishes.

Ownership of English: Widdowson (1994) and McKay (2002) argue that because English is a global language, it is no longer the sole property of "native speakers." Therefore, local varieties of English and local cultural values have a legitimate place in ELT.

Pedagogical Implications: This perspective shifts the goal of ELT from mimicking native-speaker norms to achieving intercultural communicative competence (Byram, 1997). The focus moves to intelligibility and the ability to negotiate meaning in diverse cultural contexts.

In the context of ELT, glocalization represents a paradigm shift away from linguistic imperialism and methodological dogmatism. It empowers local educators to adapt global ELT trends—such as

communicative approaches or digital literacy—specifically to their own classroom realities, valuing local languages and cultures as essential bridges to learning English.

3. Methodology

This study employs a Qualitative Document Analysis method. The primary object of study is the 9th-grade English coursebook 'Uptown', published by MoNE for the 2023-2024 academic year and beyond.

Data Analysis Procedure:

The textbook was examined unit by unit. Texts, dialogues, and visual materials were coded based on the following criteria:

Cultural Reference: (Source, Target, or International).

Value Integration: (Presence of root values like benevolence, responsibility, or respect).

Contextual Setting: (Local Turkish settings vs. Foreign settings).

4. Findings

The analysis of 'Uptown' reveals a distinct departure from previous coursebooks in its treatment of culture and values.

4.1. Predominance of Source Culture (Turkish Culture)

Unlike traditional textbooks that focus heavily on London or New York, 'Uptown' frequently utilizes Turkish settings and characters to contextualize grammar and vocabulary.

Visuals: Illustrations often depict diverse Turkish demographics and identifiable Turkish landmarks.

Reading Passages: Texts include biographies of Turkish scientists, historical figures, and athletes, allowing students to discuss their own heritage in the target language.

4.2. Integration of "Root Values"

The TCEM's emphasis on moral education is evident in the thematic choices of the units.

Benevolence and Empathy: Dialogue activities often revolve around social responsibility projects, helping the elderly, or community service, rather than purely consumerist transactions (e.g., shopping or tourism).

Family and Friendship: The unit themes prioritize strong family bonds and peer support, reflecting the collectivist nature of Turkish culture while teaching English relational vocabulary.

4.3. Intercultural Communicative Competence

While Source Culture is emphasized, 'Uptown' does not ignore the global nature of English. However, it reframes the interaction. Instead of Turkish students trying to imitate British culture, characters in the book are often portrayed explaining Turkish culture to foreigners or comparing cultural practices. This fosters Intercultural Communicative Competence—the ability to act as a mediator between cultures.

5. Discussion

The findings suggest that 'Uptown' successfully aligns with the TCEM's goal of raising "virtuous" individuals. By embedding English language instruction within the framework of local values, the textbook aims to lower the "affective filter" of students; they are learning to talk about their lives and their reality, rather than an abstract foreign concept.

This shift presents a "Instrumental" view of English: English is viewed not as a tool for assimilation into Western culture, but as a tool for narrating the Turkish experience to the world.

6. Conclusion

The Türkiye Century Education Model represents a strategic pivot in educational policy, and the 'Uptown' coursebook is a tangible manifestation of this change. This study concludes that 'Uptown' significantly differs from its predecessors by:

- Prioritizing Source Culture (Turkish) over Target Culture.
- Explicitly embedding moral and ethical values into language activities.
- Positioning English as a medium for expressing national identity.

For teachers, this implies a need to adapt their pedagogy to focus more on meaning and value-based discussions, ensuring that the language classroom becomes a space for both linguistic and moral development.

References

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.

Canagarajah, A. S. (2005). *Reclaiming the Local in Language Policy and Practice*. Lawrence Erlbaum.

Tomlinson, B. (2011). *Materials Development in Language Teaching*. Cambridge University Press.

Cortazzi, M., & Jin, L. (1999). Cultural mirrors: Materials and methods in the EFL classroom. In E. Hinkel (Ed.), *Culture in second language teaching and learning*. Cambridge University Press.

Gray, J. (2002). The global coursebook in English language teaching. Godby, M. (Ed.) *Gender, Race, and Class in Media*.

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford University Press.

Kumaravadivelu, B. (2006). *Understanding Language Teaching: From Method to Postmethod*. Lawrence Erlbaum.

McKay, S. L. (2002). Teaching English as an International Language. Oxford University Press.

Ministry of National Education (MoNE). (2023). *Türkiye Yüzyılı Maarif Modeli*. Ankara.

Robertson, R. (1995). Glocalization: Time-space and homogeneity-heterogeneity. In M. Featherstone, S. Lash & R. Robertson (Eds.), *Global Modernities*. Sage.

AUTHOR BIODATA

Ahmet Selçuk Akdemir is an associate professor of English Language Teaching (ELT) at Malatya Turgut Özal University, Turkey. Dr. Akdemir's research interests cover L2 listening, WTC, ELT pedagogy and language teacher training. As well as conducting editorial roles in several academic journals, Dr. Akdemir teaches EFL courses at undergraduate and graduate levels.

Kahramanmaraş, Türkiye